Use of prokaryotic transcriptional activators as metabolite biosensors in eukaryotic cells

Jensen, Michael Krogh; Keasling, Jay; Skjødt, Mette Louise; Snoek, Tim

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 January 2018 (11.01.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/007526 A1

(51) International Patent Classification:
C12N 15/81 (2006.01) *C12N 15/67* (2006.01)

(21) International Application Number:
PCT/EP2017/066943

(22) International Filing Date:
06 July 2017 (06.07.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16178220.6 06 July 2016 (06.07.2016) EP

(71) Applicant: **DANMARKS TEKNISKE UNIVERSITET** [DK/DK]; Anker Engelunds Vej 101 A, 2800 Kgs. Lyngby (DK).

(72) Inventors: **JENSEN, Michael, Krogh**; Hellebækgade 15, st., 2200 Copenhagen N (DK). **KEASLING, Jay, D.**; 828 Contra Costa Ave., Berkeley, California 94707-1920 (US). **SKJOEDT, Mette, Louise**; Elstedvej 23, 2610 Rødovre (DK). **SNOEK, Tim**; Sankt Hans Gade 17, 4. th., 2200 Copenhagen N (DK).

(74) Agent: **INSPICOS P/S**; Kogle Allé 2, 2970 Hørsholm (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

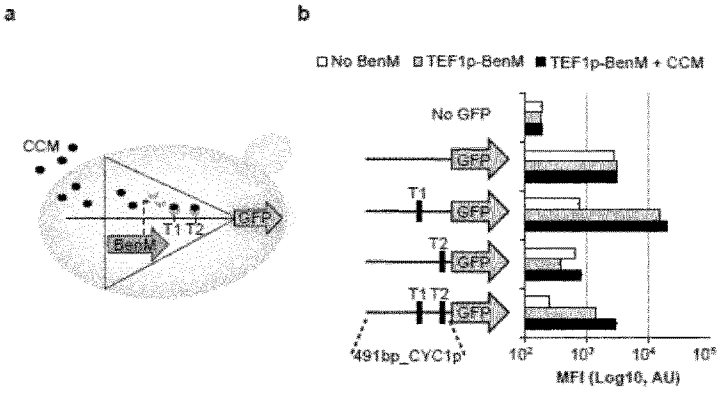
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: USE OF PROKARYOTIC TRANSCRIPTIONAL ACTIVATORS AS METABOLITE BIOSENSORS IN EUKARYOTIC CELLS

FIG. 1



(57) Abstract: The present invention relates to the use of transcriptional activators from prokaryotic organisms for use in eukaryotic cells, such as yeast as sensors of intracellular and extracellular accumulation of a ligand or metabolite specifically activating this transcriptional activator in a eukaryot, such as yeast cell, such as a cell engineered to produce this ligand. The transcriptional activator controls a promoter upstream of one or more gene, which may include e.g. a reporter gene that may be a fluorescence marker, such as luciferase, green fluorescent protein or a gnee encoding antibiotic resistance.

WO 2018/007526 A1

USE OF PROKARYOTIC TRANSCRIPTIONAL ACTIVATORS AS METABOLITE BIOSENSORS IN EUKARYOTIC CELLS

FIELD OF THE INVENTION

The present invention relates to the use of transcriptional activators from prokaryotic organisms for use in eukaryotic cells, such as yeast as sensors of intracellular and extracellular accumulation of a ligand or metabolite specifically activating this transcriptional activator in a eukaryote, such as yeast cell, such as a cell engineered to produce this ligand. The transcriptional activator controls a promoter upstream of a gene that may include e.g. a reporter gene that may be a fluorescence marker, such as luciferase or green fluorescent protein.

BACKGROUND OF THE INVENTION

Whole-cell biocatalysts have proven a tractable path towards sustainable production of bulk and fine chemicals. Yet, screening libraries of cellular designs to identify best-performing biocatalysts is most often a low-throughput endeavour. For this reason the development of biosensors enabling real-time monitoring of product formation has gained significant attention.

Bio-based production of chemicals and fuels is an attractive avenue to reduce dependence on petroleum. For bio-based production, biocatalysts must often be genetically modified in order to increase product titers, rates and yields. However, the current efficiency of genome engineering methods and parts prospecting allow for unprecedented genotype diversity that vastly outstrips our ability to screen for best cell performance.

To meet this demand, bioengineers have started to develop genetically encoded devices and systems that enable screening and selection of better-performing biocatalysts in higher throughput. Genetic devices like oscillators, amplifiers and recorders, which have been developed based on fine-tuned relationships between input and output signals are promising tools for programming and controlling gene expression in living cells. These devices sense extra- or intracellular perturbations and actuate cellular decision-making processes akin to logic gates in electrical circuits. For instance, AND logic gates have been built using inducible expression of two split T7 RNA polymerase domains which control reporter gene expression

when used in combination with a complimentary set of altered T7-specific promoters. From a diverse set of inputs, other molecular gating components like RNA aptamers, allosteric regulators and ligand-binding transcription factors have been engineered to control outputs for applications such as high-throughput screening, actuation on cellular metabolism, and evolution-based selection of optimal cell performance.

A key component in many of the reported devices is a ligand-inducible transcriptional regulator. Transcriptional regulators are powerful components finding many uses in genetic designs. Owing to their modular structure, transcriptional regulators have proven to be versatile platforms for genetically encoded Boolean logic functions. In particular, gene switches based on ligand-binding transcriptional repressors bind to genomic targets in the absence of their cognate ligand and thereby repress gene expression of the downstream gene(s), whereas binding between ligand and repressor causes the release of the repressor from the DNA and thereby a de-repression. In NOT gates like this, the simple steric hindrance of RNA polymerase progression, like in the case of the tetracyclin-responsive gene switch TetR, have for decades been used for conditional control of gene expression in both prokaryotic and eukaryotic chassis. Most importantly, transcriptional repressors and other artificial transcriptional regulators can be further engineered - including the addition of nuclear localization signals, destabilization domains and transcriptional activation regions - to repurpose conditional repressors into activators. Though conceptually intriguing and practically relevant, the repurposing and engineering of logic gates can suffer from the inherent need for extensive engineering.

Though most ligand-inducible genetic devices adopted for eukaryotes have historically been founded on transcriptional repressors, a hitherto untapped resource for use in genetic designs is ligand-inducible transcriptional activators. Remarkably, bacterial genomes encode a multitude of ligand-inducible activators amenable for integration into synthetic genetic devices. In bacteria, transcriptional activation takes place through (i) a promoter-centric or (ii) an RNA polymerase-centric mechanism. In the former case a transcriptional activator can bind to an operator site in a promoter thereby improving its ability to guide RNA polymerase to initiate transcription, whereas in the latter case activation relies on interactions with the RNA-polymerase itself such as when a housekeeping σ factor is replaced by another σ factor. Examples of prokaryotic transcriptional activators used for genetic designs in other non-native prokaryotic chassis include arabinose-inducible AraC and quorum sensing LuxR. However, so far no direct transplantation of prokaryotic ligand inducible transcriptional activators has been reported in eukaryotes. As the number of bulk and fine chemicals produced in eukaryote chassis continues to increase, there is an increasing need to be able to

regulate these pathways using any and all available means, including the heretofore untapped prokaryotic ligand-inducible transcriptional activators.

OBJECT OF THE INVENTION

It is an object of embodiments of the invention to provide eukaryotic cells comprising bacterial transcriptional activator systems functioning in a eukaryotic chassis. Accordingly some genes within the eukaryotic cell are under influence of a bacterial transcriptional activator preferably working in a eukaryotic cell on the endogenous promoter of this eukaryotic cell.

It is an object of embodiments of the invention to provide eukaryotic cells that contain sensors, such as easily visible sensors of intracellular and extracellular accumulation of a ligand or metabolites being produced by this cell.

The inventors of the present invention have applied systematic engineering of multiple parameters to search for a general biosensor design based on small-molecule binding transcriptional activators from the prokaryote super-family of LysR-type transcriptional regulators (LTTRs). The present inventors have identified a design supporting LTTR-dependent activation of reporter gene expression in the presence of cognate small-molecule inducers. As proofs-of principle they have applied the biosensors for in vivo screening of cells producing naringenin or cis, cis-muconic acid at different levels, and show that reporter gene output correlates with product accumulation. The transplantation of prokaryotic transcriptional activators into a eukaryotic chassis illustrates the potential of a hitherto untapped resource for engineering biosensors useful for biotechnological applications.

The present inventors have found a direct onboarding of a prokaryotic transcriptional activator as a biosensor for e.g. cis, cis-muconic acid (CCM) in budding yeast Saccharomyces cerevisiae. Based on a multi-parametric engineering strategy the present inventors identified a functional design for the biosensor. Most importantly the design is applicable to a range of other LTTR-based biosensors founded on small-molecule induced transcriptional activators. As proofs-of-principle two of these biosensors were applied for real-time monitoring of bulk and fine chemical product accumulation in yeast cells engineered to produce CCM and naringenin, respectively. This constitutes the first successful direct transfer of prokaryotic transcriptional activators into a eukaryotic chassis to activate gene expression merely by placing the binding site of a transcriptional activator at a defined location in a reporter promoter and without reconfiguring any other motifs and domains.

Systematic engineering and meticulous characterization have for decades pushed forward the sequence-function understanding of genetic parts and interactions thereof. This has allowed the rational engineering of parts and genetic circuits useful for a range of applications within biotechnology. While most of the genetic devices originate from prokaryotes, transplantation into eukaryotes has been reported for a number bioswitches in order to construct orthogonal genetic devices to control a cellular response to a defined input. Specifically, genetic devices enabling the manipulation of transcription through the transplantation of prokaryote transcriptional repressors have inspired researchers, in their quest for tools to screen, select and actuate on cellular responses. In this study we have shown that ligand-inducible transcriptional activators from the largest family of transcriptional regulators found in prokaryotes, can be ported to eukaryotic chassis and used to measure the level of a small molecule inside the cell and activate transcription. The LTTR-based transcriptional activators function as is in yeast without any further engineering nor the co-expression of other molecular components (i.e. σ factors). In fact, through a systematic engineering approach we provide a framework from which new ligand-binding transcriptional activators from the LTTR family can be designed through the simple swapping of a candidate LTTR operator sequence into the 209bp_CYC1p truncated endogenous promoter at a defined position (T1)(Fig. 2a-b, Table 4, and Table 5). Also, and most importantly, we provide two successful proofs-of-principle for such biosensors to screen *in vivo* for the best-performing biocatalysts.

Compared to many of the studies using transcriptional repressors as biosensors in eukaryotes, the biosensor outputs based on ligand-inducible transcriptional activators presented in this study have lower dynamic ranges falling within one order of magnitude. This is in agreement with the observation from using BenM and FdeR as biosensors in *E. coli*. This can pose a challenge for their applicability in genetic designs where a larger dynamic range is needed. However, we demonstrated in this study how these biosensors could be subjected to biosensor-based FACS for identification of biosensor designs with improved characteristics (ie. dynamic range), which may expand their applicability for metabolic engineering. For sure, we envision this to be exploitable for high-throughput screening of libraries of genetic designs for metabolites for which there exists no high-throughput screening assay or biosensor.

Apart from dynamic range, another key performance measure for biosensors, is their operational range. In our study we demonstrated how biosensors could be used in laboratory strains with limited engineering towards titer improvements, which at their best still are far from commercially relevant. Indeed, in diploid yeast, production of 559.3 mg/L CCM was recently reported, whereas an *E. coli-E. coli* co-cultivation study have reported the production of 2 g/L CCM. Though tolerance to low-pH fermentations should make yeast an economically feasible chassis for biobased production of dicarboxylic acids like CCM, the CCM biosensor

design based on BenM may need to be adjusted or evolved as production hosts become better and the product titers gets higher. Additionally, the biosensor will need to be matched to the production kinetics of the individual strain or library of biocatalysts.

Nevertheless, the LTTR-based ligand-inducible transcriptional activators reported here are much-needed tools for optimizing biocatalysts that produce chemicals and fuels for which there exist no high-throughput screen or selection. This should spur interest in developing many other orthogonal logic gates based on LTTR members, which could serve as a vast and valuable reservoir for developing new ligand-inducible genetic circuits capable of high-throughput screening, reprogramming and growth-coupled strain selection for bio-based production of chemicals. Furthermore, as the mode-of-action of transcriptional activators (YES) differ from that of repressors (NOT), the future possibility for higher-order designs within cellular reprogramming can now be exploited in greater diversity.

SUMMARY OF THE INVENTION

It has been found by the present inventor(s) that transcriptional activators from prokaryotic organisms may work in a eukaryotic chassis by the positioning of the operator at a particular place within the eukaryotic promoter. The inventors found that this can be used e.g. for providing biosensors for intracellular and extracellular accumulation of a ligand or metabolites produced within a eukaryotic cell.

So, in a first aspect the present invention relates to a eukaryotic cell, such as a yeast cell comprising a bacterial transcriptional activator and a corresponding operator sequence positioned in a eukaryotic promoter, such as positioned within an endogenous promoter of said cell, which activator controls the expression of a gene from said eukaryotic promoter.

In a second aspect the present invention relates to a eukaryotic cell, such as a yeast cell comprising a bacterial transcriptional activator and a corresponding operator sequence, which activator controls the expression from a eukaryotic promoter, such as an endogenous promoter of the cell in response to a ligand specifically binding the transcriptional activator.

In a third aspect the present invention relates to a eukaryotic cell, such as a yeast cell comprising a bacterial transcriptional activator and a corresponding operator sequence positioned in a eukaryotic promoter, such as positioned within an endogenous promoter of the cell, which activator controls the expression of a gene from the eukaryotic promoter depending on the presence, such as dose dependent, of a ligand specifically binding the transcriptional activator

In a further aspect the present invention relates to the use of a prokaryotic transcriptional activator as a regulator of transcription in a eukaryotic cell, such as a yeast cell according to the invention; the transcriptional activator being activated by a ligand specifically binding the transcriptional activator to induce the expression of a protein product from a eukaryotic promoter of the cell; the promoter containing the operator sequence corresponding to the transcriptional activator.

In a further aspect the present invention relates to a the use of a prokaryotic transcriptional activator as a regulator of transcription in a eukaryotic cell, such as a yeast cell according to the invention; the transcriptional activator being activated depending on the presence, such as dose dependent, of a ligand specifically binding the transcriptional activator to induce the expression of a protein product from a eukaryotic promoter of said cell, the promoter containing the operator sequence corresponding to said transcriptional activator positioned within the promoter.

In a further aspect the present invention relates to a the use of a prokaryotic transcriptional activator as a metabolite biosensor for measuring the amount of a ligand extracellular of and/or produced by a eukaryotic cell, such as a yeast cell according to the invention, wherein the ligand specifically bind the transcriptional activator to induce expression of a reporter gene from a eukaryotic promoter of the cell, the promoter containing the operator sequence corresponding to the transcriptional activator.

In a further aspect the present invention relates to a method for measuring the amount of a ligand intracellular or extracellular of a eukaryotic cell, such as a yeast cell; said cell comprising a bacterial transcriptional activator and a corresponding operator sequence, which activator controls the expression of a reporter gene from a eukaryotic promoter of said cell in response to said ligand specifically binding said transcriptional activator; said promoter containing the operator sequence corresponding to said transcriptional activator; said method including the steps of

a) Cultivating a eukaryotic cell according to the invention;

b) Measuring the output from said promoter of said reporter gene;

c) Correlating said output from step b) with amount of said ligand.

In some embodiments the ligand is not produced by the eukaryotic cell, but is present in a solution of the cultivation medium of the eukaryotic cell, or such as when used to report toxic waste in a soil.

In a further aspect, the present invention relates to a recombinant transcriptional activator with increased activity, such as BenM with mutations at any one or more of the positions H110R, F211V, and Y286N.

LEGENDS TO THE FIGURE

Figure 1. Onboarding the *cis-cis*-muconic acid (CCM) responsive prokaryotic transcriptional activator BenM in yeast. (a) Schematic outline of native and synthetic full-length (491bp) CYC1 promoter variants with different *BenO* positioning and number (T1 and/or T2). The transcriptional activator BenM from *Acinetobacter* sp. ADP1 controls expression of GFP from the synthetic CYC1 promoter with BenM operator (*BenO*) integrated at position T1 and/or T2. CCM further induces BenM-dependent expression of GFP (b) Mean fluorescence intensity (MFI) values from flow cytometry measurements of GFP intensities in the presence or absence of BenM expressed from the constitutively active TEF1 promoter, and following 24 h of incubation in the presence or absence of 1.4 mM CCM. (c) Screening 84 yeast strains expressing all possible combinations of BenM expression levels (TDH3p, TEFp, RNR2p and REV1p) individually or in combination with native or engineered CYC1p reporter promoters of different lengths (491bp, 272bp, 249bp and 209bp), *BenO* positioning, and number (T1 and/or T2) by flow cytometry. Outputs are ordered according to GFP intensity in control medium. Dashed lines indicate background fluorescence as inferred from BenM expressing strains without GFP, and arrow indicate best-performing biosensor design. Genotypes and GFP expression levels of all 84 strains are listed in Tables 1 and 2, respectively. (d) Heat-maps showing fold change (FC) in GFP expression between CCM-induced and control cultures of 80 strains shown in (c). For (b) and (c) mean fluorescence intensities (MFI) are shown as mean ± s.d. from three (n=3) biological replicate experiments. AU, arbitrary units.

Figure 2. High-throughput engineering and screening of BenM variants with improved CCM-inducibility. (a) Purified products from three rounds of error-prone PCR (epPCR) using the effector-binding domain (EBD) of BenM as template, were co-transformed into yeast together with a linearized centromeric plasmid, to allow for *in vivo* library reconstitution by gap repair and expression of wild-type BenM DNA binding domain (DBD) fused to approx. 40,000 variants of the EBD. Transformed yeast contained a chromosomal integration of 209bp_CYC1p_*BenO*_T1 controlling the expression of GFP to allow for FACS-based screening of BenM variants with improved CCM-inducibility. (b) Representative flow cytometry histograms of fluorescence intensities obtained from a population of yeast cells expressing CCM sensor variants in control (grey) and CCM-inducer (light green) media. Control, CCM-induced and sorted (darker green) cell populations are normalized to mode for comparison. The proportions of cells within each histogram were calculated by FlowJo software as

described in Methods (c) Isolated BenM variants were grown as clonal cultures, validated by flow cytometry and the EBDs of variants with significantly higher GFP expression under CCM-induced cultivation were sequenced. Mean fluorescence intensities (MFI) are shown as mean ± s.d. from three (*n*=3) biological replicate experiments. AU, arbitrary units. (d) Ribbon representations of the EBD of BenM (PDB 2F7A) with the residue changes identified in $BenM^{H110R, F211V, Y286N}$ highlighted in green. Bound CCM is highlighted as a magenta Van der Waals sphere.

Figure 3. Biosensor specificity and transcriptional orthogonality. (a) The specificity of the CCM biosensor was tested by addition of various dicarboxylic acids (1.4 mM) to the growth medium. GFP expression was measured by flow cytometry following 24 h of cultivation. (b) RNA sequencing FPKM (fragments per kilo base per million) are plotted for yeast cells stably expressing 209bp_CYC1p_*BenO*_T1::GFP reporter construct and $BenM^{H110R, F211V, Y286N}$ versus cells only expressing the reporter construct following 24 h of cultivation in medium supplemented with CCM. Purple area indicates 2-fold cut-off and red dots significantly differentially regulated genes as inferred from cuffdiff (>2-fold, $P<0.05$)(see also Fig. 5). All data points are averaged from three (*n*=3) biological replicates.

Figure 4. Onboarding transcriptional activators from the LTTR family as biosensors in yeast. (a) *Left:* Schematic illustration of LTTR-mediated activation of GFP expression by binding to the cognate operator in position T1 of 209bp_CYC1p. *Right:* The 209bp_CYC1p_T1 reporter promoter design supports GFP expression when controlled by individual LTTR transcriptional activators expressed from either a weak (REV1p) or a strong (TDH3p) promoter. The y-axis shows fold induction in mean fluorescence intensity (MFI) in cells expressing individual LTTRs relative to cells not expressing the LTTRs. (b) *Left:* Schematic illustration on external application of individual ligands for induction of LTTR-mediated activation of GFP expression. *Right:* External application of individual ligands can induce LTTR-mediated activation of GFP expression. The y-axis shows fold induction in mean fluorescence intensity (MFI) for cells grown for 24 h in medium containing either *cis, cis*-muconic acid (CCM), naringenin (NAR), L-arginine (ARG), protocatechuic acid (PCA) or malonic acid (MAL) compared to cells growing in control medium. (c) Heatmap showing orthogonality of MdcR- and ArgP-mediated transcriptional regulation of GFP expression controlled by either reporter promoter 209bp_CYC1p_Mdc*O*_T1 or 209bp_CYC1p_Arg*O*_T1 (Table 4). Color key shows mean fluorescence intensity (MFI) from three (*n*=3) biological replicate experiments. For (a) and (b), mean fluorescence intensity (MFI) values and their error bars are calculated as mean ± s.d. from three (*n*=3) biological replicates.

Figure 5. Biosensor sensitivity and operational range. (a) The response functions of wild-type and engineered BenM$^{H110R, F211V, Y286N}$ expressed in yeast from REV1p as measured by flow cytometry using various concentrations of CCM (24 h) and the 209bp_CYC1p_*BenO*_T1 promoter controlling the expression of GFP. A yeast strain without BenM expressed is used as a control for background GFP fluorescence from the 209bp_CYC1p_*BenO*_T1 promoter. (b) The response function measurement for the naringenin biosensor when FdeR is expressed from a weak (REV1p) or a strong (TDH3p) promoter using various concentrations of naringenin (24 h) and the 209bp_CYC1p_*FdeO*_T1 reporter promoter controlling the expression of GFP. A yeast strain without FdeR expressed is used as a control for background GFP fluorescence from the 209bp_CYC1p_*FdeO*_T1 promoter. For (a) and (b) mean fluorescence intensity (MFI) values and their error bars are calculated as mean ± s.d. from three (n=3) biological replicate experiments

Figure 6. *In vivo* application of CCM and naringenin biosensors in yeast. (a) Schematic representation of the heterologous 3-step CCM production pathway for testing BenM as a biosensor for *in vivo* CCM production in yeast. Additionally, over-expression of Tkl1 was included together with balancing of the heterologous three-step pathway (*Pa*AroZ, *Kp*AroY and *Ca*CatA) using single or multi-loci integration of AroY subunits B and C (Iso, isoform)(see Methods and Table 1). (b) Following 24 h of cultivation, CCM titers and MFIs were quantified and plotted for each strain. (c) Schematic representation of heterologous 5-step naringenin production pathway adopted from Naesby *et al.* For the hydroxylation of cinnamate to coumarate a fusion protein of *At*C4H and *At*ATR2 was used. For testing FdeR as a biosensor for *in vivo* naringenin production in yeast, mean fluorescence intensity (MFI) in three different strains engineered with one copy of the 5-step naringenin production pathway (EVR1) or with one (EVR2) or two (EVR3) additional integrations of bottleneck enzymes were compared to a control strain (EVR0, ctrl) without the production pathway. Following 48 h of cultivation, naringenin titers and MFIs were quantified and plotted. For both (b) and (d) data are average of three biological replicates. Mean fluorescence intensity (MFI) values and metabolite quantifications are presented as means ± s.d. from three (n=3) biological replicate experiments.

Fig. 7 BenM regulation of the ben operon during benzoate catabolism in Acinetobacter sp. ADP1 is feed-back induced by the intermediate catabolite cis,cis-muconic acid (CCM). Upon detection of CCM from the aromatic acid catabolism, the constitutively DNA-bound BenM tetramer undergoes a conformational change facilitating the accessibility of RNA polymerase and active transcription of the ben operon (Bundy, B. M., Proc. Natl. Acad. Sci. U. S. A. 99, 7693–8 (2002). Sequence of the BenM operator (BenO). The three potential binding sites for BenM are highlighted in blue with site 1 displaying dyad symmetry exactly matching the

consensus sequence of LysR-type regulators (Collier, L. S., J. Bacteriol. 180, 2493–501 (1998)).

Fig. 8 (a) Sequence outline of the full-length CYC1 promoter with native upstream activating sequences (UASs) shown in blue and operator sites for TATA-binding proteins shown in red. Sites for positioning of BenM operators are marked with black triangles (T1 and T2) and sites for truncations (272bp, 249bp and 209bp) marked with dashed vertical lines. USER cloning site and Kozak sequence is italicized upstream the open reading frame of yeast-enhanced GFP (bold). (b) Sequence of the FdeR, PcaQ, ArgP and McdR operators used to swap into position T1 of the 209bp_CYC1p shown in (a) (Siedler, S., Metab. Eng. 21, 2–8 (2014) and Maclean, A. M., Microbiology 157, 2522–33 (2011).

Fig. 9 (a) Uptake of cis,cis-muconic acid (CCM) at pH 4.5 by S. cerevisiae cells following 24 h of growth. (b) Representative growth curves of yeast cells in liquid medium containing different concentrations of CCM. OD values were determined at 1-h intervals over 25-h period. For both (a) and (b) data display means ± s.d. from three (n=3) biological replicate cultivations.

Fig. 10 Screening 84 yeast strains expressing all possible combinations of BenM expression levels (TDH3p, TEFp, RNR2p and REV1p) individually or in combination with native or engineered CYC1p reporter promoters of different lengths (491bp, 272bp, 249bp and 209bp), BenO positioning, and number (T1 and/or T2) by flow cytometry after 24 h of growth in control medium or medium supplemented with 1.4 mM CCM. Outputs are ordered according to GFP intensity in control medium. Dashed lines indicate background fluorescence as inferred from strains expressing only BenM (no reporter). Genotypes and GFP expression levels of all 84 strains can be found in Tables 1 and 2, respectively. Mean fluorescence intensity (MFI) values and their error bars are calculated as mean ± s.d. from three (n=3) biological replicate experiments. AU, arbitrary units.

Fig. 11 (a) A box and whisker plot showing the mean value, 1 and 99 percentiles for three (n=3) biological replicate RNA sequencing experiments. Outliers are depicted as black dots. The GFP is highlighted in green. RNA was collected following 24 h of cultivation in mineral medium pH 4.5 with 1.4 mM CCM. The fold change in FPKM is displayed as a log2 normalized value for all expressed genes. (b) A fold change histogram of FPKM showing the fold changes (log2) in gene expression from Fig. 2c plotting strain including BenMH110R, F211V, Y286N (MeLS0284) over the strain without BenMH110R, F211V, Y286N (MeLS0138). GFP is indicated with an arrow. The data are representative of three (n=3) biological replicates.

Fig. 12 Endogenous response function of the CCM and naringenin biosensors. (a) Cultivation medium was analyzed for CCM concentration by LC-MS and flow cytometry performed for GFP intensity measurements of six different CCM producing strains compared to a reference CCM null background strain (see Table 1) following 24 h and 72 h cultivations. (b) Average titers for the six CCM-producing strains at 24 h and 72 h of cultivation, compared to the reference strain. (c) Cultivation medium was analyzed for naringenin concentration by UPLC and flow cytometry performed for GFP intensity measurements of three different naringenin producing strains compared to a reference naringenin null background strain following 24 and 48 h cultivations. (d) Average titers for the three naringenin-producing strains at 24 h and 48 h of cultivation, compared to the reference strain. For (a-d) data are presented as means ±s.d. from three (n=3) biological replicate experiments. Table 1 lists all strain genotypes.

Figure 13. BenM activates reporter expression in CHO. CHO cells were transfected with a plasmid with the BenO-containing human cytomegalovirus (CMV) promoter controlling the expression of GFP as well as an empty vector (- BenM) or a vector expressing BenM (+ BenM). Total GFP expression was measured after 24 h, and normalized by total RFP expression. Average and standard deviation are based on three biological replicates. *; $p<0.05$ (t-test).

Fig. 14. Screening 17 yeast strains expressing BenM from the REV1p in combination with CYC1p reporter promoter of 209bp with BenO placed at different positions upstream of TATA1. The fold induction (mean ± s.d.) was calculated by dividing mean fluorescence intensity (MFI) in medium with 1.4 mM CCM by the MFI in control medium as measured by flow cytometry after 24 h of growth for three biological replicates.

DETAILED DISCLOSURE OF THE INVENTION

Definitions

The term "eukaryotic cell" is used herein in its normal sense. The term includes any animal, mammalian, fungi, yeast, insect and algae cell. In some specific embodiments the eukaryotic cell is a yeast cell.

The term "yeast cell" refers to the single-celled microorganisms classified as members of the fungus kingdom. The term includes but is not limited to cells of a genus selected from the group consisting of Kluyveromyces, Saccharomyces and Hanensula, such as a yeast cell selected from the group consisting of Saccharomyces cerevisiae and Saccharomyces boulardii.

The term "bacterial transcriptional activator" as used herein refers to a protein, such as any know protein naturally derived from a bacterium (a transcription factor) that increases gene transcription of a gene or set of genes in this bacterium. It is to be understood that a bacterial transcriptional activator and its corresponding operator sequence (and whether it is derived from a prokaryote genome and not found in a eukaryotes motif) will be easily identified by a simple sequence search for the person skilled in the art. Essentially, for a transcriptional regulator to be defined as a prokaryotic transcriptional activators by the person skilled in the art it must adhere to all the following points:

a) The gene encoding the protein sequence is found natively in a prokaryotic genome.

b) When using the Basic Local Alignment Search Tool (BLAST; https://blast.ncbi.nlm.nih.gov/Blast.cgi) for comparing nucleotide or protein sequences to sequence databases, the query sequence aligns more to sequences of prokaryote origin in terms evolutionary relationships than to sequences of eukaryote origin.

c) In its native context of a prokaryotic genome, deletion of the gene encoding the protein sequence of the transcriptional regulator will cause lower or no change in expression of its target gene.

d) In its native context of a prokaryotic genome, over-expression of the gene encoding the protein sequence of the transcriptional regulator will not cause lower expression of its target gene, as is the case for transcriptional repressors.

e) The gene encoding the protein sequence is categorized functionally as an activator in the RegPrecise database (http://regprecise.lbl.gov) ; a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes.

In some embodiments the bacterial transcriptional activator is within the prokaryote super-family of LysR-type transcriptional regulators (LTTRs).

In some embodiment the bacterial transcriptional activator as used herein is selected from the list of tables 6 and 7.

In some embodiments the term refers to an intact transcriptional activator containing both an activation domain and a DNA-binding domain.

The term "corresponding operator sequence" as used herein refers to the DNA sequence that binds a specific bacterial transcriptional activator in order to make the bacterial transcriptional activator take effect. The operator sequence is placed within the promoter on which the activator works. The corresponding operator sequence is identified as part of the DNA sequence of a prokaryote promoter sequence, which is under the regulation of the bacterial transcriptional activator.

The term "eukaryotic promoter" as used herein refers to a region of DNA derived from or within a eukaryotic cell that initiates transcription of a particular gene downstream of this promoter.

The term "endogenous promoter" as used herein refers to a promoter that normally is present in the cell in use.

The term "ligand specifically binding a transcriptional activator" as used herein refers to a ligand, which specifically binds to a particular transcriptional activator to control the functioning of the activator in a system of a so-called ligand-inducible transcriptional regulator.

The term "exogenous" refers to a gene that originates outside of the organism of the specific cell being used.

In some embodiments, the eukaryotic cell contains a reporter gene, preferably in operative linkage with the eukaryotic promotor responsive to the bacterial transcriptional activator. Exemplary reporter genes include enzymes, such as luciferase, phosphatase, or p-galactosidase which can produce a spectrometrically active label, e. g., changes in color, fluorescence or luminescence. In some embodiments the reporter gene encodes a gene product selected from the group consisting of luciferase, green fluorescent protein, p-lactamase chloramphenicol acetyl transferase, ss- galactosidase, secreted alkaline phosphatase, p-lactamase, p-glucuronidase, alkaline phosphatase, blue fluorescent protein, and chloramphenicol acetyl transferase.

"Upstream activating sequences" as used herein refers to cis-acting elements of a eukaryotic promoter that modulate the rate of initiation of transcription well known to the person skilled in the art. Specific sequence and number of subsites or regions is specific for the promotor being used.

Specific embodiments of the invention

As described above the present invention relates to a eukaryotic cell, such as a yeast cell comprising a ligand-binding bacterial transcriptional activator and a corresponding operator sequence positioned in a eukaryotic promoter, such as positioned within an endogenous promoter of the cell, which activator controls the expression of a gene from the eukaryotic promoter.

In some embodiments the expression of a gene from said eukaryotic promoter is depending on the presence, such as dose dependent, of a ligand specifically binding the transcriptional activator.

In some embodiments the cell comprises a gene encoding the expression of the ligand, one or more genes encoding a pathway of enzymes synthesizing the ligand, and/or a gene encoding a compound that is metabolized into the ligand. In some embodiment such a gene is expressed from the eukaryotic promoter.

In some embodiments the cell comprises an exogenous reporter gene, and/or one or more further regulatory gene, such as a gene encoding antibiotic resistance. In some embodiment such a gene is expressed from the eukaryotic promoter.

In some embodiments the reporter gene provides for fluorescence output, such as a gene encoding green fluorescent protein, blue fluorescent protein or luciferase.

In some embodiments the one or more the genes independently selected from the gene encoding the expression of the ligand, one or more genes encoding a pathway of enzymes synthesizing the ligand, a gene encoding a compound that is metabolized into the ligand, an exogenous reporter gene, and one or more further regulatory gene; is under the control and/or is activated by the eukaryotic promoter.

In some embodiments the transcriptional activator is selected from any one selected from table 6, such as any one selected from BenM, FdeR, MdcR, and ArgP.

In some embodiments the ligand and transcriptional activator is selected from muconic acid and BenM; Naringenin and FdeR; Malonate and MdcR, and L-arginin and ArgP.

In some embodiments the cell is a yeast cell, such as Saccharomyces cerevisiae.

In some embodiments the cell is a mammalian cell, such as a Chinese hamster ovary cell.

In some embodiments the promoter is a full length promoter, or a truncated version with upstream activating sequences, such as UAS1 and UAS2 of the CYC promoter, removed.

In some embodiments the promoter is a yeast promoter, such as the full length CYC1 promoter or CYC1 with upstream activating sequences (UAS1 and UAS2) removed.

In some embodiments the promoter is a mammalian promoter, such as the full length CMV promoter.

In some embodiments the transcriptional activator work through a promotor-centric mechanism, wherein the transcriptional activator bind to an operator site in the promotor thereby improving its ability to guide RNA polymerase to initiate transcription.

In some embodiments the transcriptional activator does not require binding to any other regulatory subunits and/or which cell is without any further engineering or the co-expression of other molecular components regulating the transcriptional activator.

In some embodiments the transcriptional activator does not require binding to any other regulatory subunits apart from its specific ligand and/or which cell is without any further engineering or the co-expression of other molecular components regulating said transcriptional activator.

In some embodiments the operator sequence is specific for the transcriptional activator within the promoter.

In some embodiments the operator sequence is positioned immediately upstream of the TATA box, such as a TATA box 1, such as TATA-1ß, such as anywhere between 6-15 bp, such as anywhere between 6-14 bp, such as anywhere between 6-13 bp, such as anywhere between 6-12 bp, such as anywhere between 6-11 bp, such as anywhere between 6-10 bp, such as anywhere between 6-9 bp, such as anywhere between 6-8 bp, such as anywhere between 6-7 bp, such as 6 bp upstream of said TATA box of said eukaryotic promoter..

In some embodiments the operator sequence is positioned immediately upstream of one of the two TATA boxes - TATA-1ß, such as anywhere between 6-15 bp upstream of TATA box 1, such as anywhere between 6-14 bp upstream of TATA box 1, such as anywhere between 6-13 bp upstream of TATA box 1, such as anywhere between 6-12 bp upstream of TATA box 1, such as anywhere between 6-11 bp upstream of TATA box 1, such as anywhere between 6-10 bp upstream of TATA box 1, such as anywhere between 6-9 bp upstream of TATA box 1, such as anywhere between 6-8 bp upstream of TATA box 1, such as anywhere between 6-7 bp upstream of TATA box 1, such as 6 bp upstream of TATA box 1.

In some embodiments the operator sequence is positioned immediately 6 bp upstream of the TATA box - TATA-1ß, such as anywhere between 6-15 bp upstream of TATA box 1, such as anywhere between 6-14 bp upstream of TATA box 1, such as anywhere between 6-13 bp upstream of TATA box 1, such as anywhere between 6-12 bp upstream of TATA box 1, such

as anywhere between 6-11 bp upstream of TATA box 1, such as anywhere between 6-10 bp upstream of TATA box 1, such as anywhere between 6-9 bp upstream of TATA box 1, such as anywhere between 6-8 bp upstream of TATA box 1, such as anywhere between 6-7 bp upstream of TATA box 1, such as 6 bp upstream of TATA box 1.

In some embodiments the transcriptional activator belongs to the prokaryote super-family of LysR-type transcriptional regulators (LTTRs).

In some embodiments the operator is an LTTR operator sequence selected from *BenO, FdeO, MdcO*, and *ArgO*.

In some embodiments the transcriptional activator is co-expressed in the cell, such as from a promoter selected from TEF1, REV1, RNR2 and TDH3.

In some embodiments the transcriptional activator is a functional variant with increased activity, such as BenMH110R, F211V, Y286N.

EXAMPLE 1

Strains, chemicals and media. *Saccharomyces cerevisiae* CEN.PK102-5B (MAT*a ura3-52 his3Δ1 leu2-3/112 MAL2-8^c SUC2*), CEN.PK113-5A (MAT*a*, *trp1 his3Δ1 leu2-3/112 MAL2-8^c SUC2)* and CEN.PK113-7D (wild type, *MAT*a *MAL2-8^c SUC2)* strains were obtained from Peter Kötter (Johann Wolfgang Goethe-University Frankfurt, Germany). In principal any other yeast strains may be used, such as one obtained from public repository EuroScarf. EasyClone plasmids used in this work are described in Jensen, N. B. et al. EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. 14, 238–48 (2014). *Escherichia coli* strain DH5α was used as a host for cloning and plasmid propagation. The chemicals and Pfu TURBO DNA polymerase were commercially obtained (Sigma-Aldrich and Agilent Technologies Inc., respectively). All acids used were >97% in purity. *S. cerevisiae* cells were grown at 30°C in synthetic complete medium as well as drop-out media and agar plates were prepared using pre-mixed drop-out powders (Sigma-Aldrich). Mineral medium was freshly prepared as described previously. For all media containing diacids, 1.4 mM of the individual diacids were dissolved in mineral medium and pH adjusted to 4.5 before sterile filtration. For CCM several dilutions were made to examine the performance of the CCM biosensor. For naringenin, mineral medium was supplemented with 0, 0.05, 0.10 or 0.20 mM naringenin, dissolved in ethanol, the final ethanol concentration for each medium was 2% (v/v), and the final pH of the medium was adjusted to 6.0. *E. coli* cells were grown at 37°C in Luria-Bertani (LB) medium supplemented with 100 μg/mL ampicillin.

Synthetic genes and oligonucleotides. Oligonucleotides and synthetic genes were commercially synthesized (Integrated DNA Technologies, Inc. and Thermo Fisher Scientific Inc., respectively). Sequences of synthetic genes and oligonucleotides can be found in Tables 4 and 5, respectively.

Plasmids, strains and library construction. Except *Arabidopsis thaliana At4CL-2* (NM_113019.3) and *Saccharomyces cerevisiae ScTKL1* (NM_001184171.1), all genes encoding *Klebsiella pneumoniae AroY.B* (AAY57854.1), *AroY.Ciso* (BAH20873.1), *AroY.D* (AAY57856.1), *Candida albicans CatA* (XP_722784.1), *Podospora anserina AroZ* (XP_001905369.1), *Acinetobacter* sp ADP *BenM* (AAC46441.1), *Arabidospis thaliana AtC4H* (NM_128601.2), *Arabidopsis thaliana* AtATR2 (NM_179141.2), *Arabidopsis thaliana AtPAL2* (NM_115186.3), *Petunia hybrida PhCHI (X14589)*, *Hypericum androsaemum HaCHS* (AF315345), *Schizosaccharomyces pombe MAE1* (NM_001020205.2), *Sinorhizobium meliloti PcaQ* (NC_003078.1), *Escherichia coli ArgP* (NC_000913.3), *Klebsiella pneumonia MdcR* (U14004), and *Herbaspirillum seropedicae* SmR1 *FdeR* (Hsero_1002, UniProtKB - D8J0W4_HERSS) were codon-optimized for expression in yeast (see Table 4 for full sequences). All gene fragments and correct overhangs for USER-cloning were amplified by PCR using oligonucleotides listed and described in Table 5. Unless otherwise stated the amplified products were USER cloned into EasyClone integrative plasmids Jensen, N. B. et al. (2014), and confirmed by sequencing.

The list of the constructed plasmids can be found in Table 3. Transformation of yeast cells was carried out by the lithium acetate method Gietz, R. D. & Schiestl, R. H. Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2, 38–41 (2007), and strains selected on synthetic drop-out medium (Sigma-Aldrich), selecting for appropriate markers. For selection of strain carrying KanMX and HypMX, the media was supplemented with 200 μg/mL G418 sulphate and 200 μg/mL hygromycin B, respectively. Transformants were genotyped using oligonucleotides described in Table 5. The resulting strains are listed in Table 1.

To establish the CCM producing strains, we expressed the dehydroshikimate DHS dehydratase from *P. anserina* (*Pa*AroZ), the PCA decarboxylase genes from *K. pneumoniae* (*Kp*AroY.B, *Kp*AroY.Ciso, *Kp*AroY.D), and the catechol 1, 2 dioxygenase CDO from *Candida albicans* (*CaCatA*) in *S. cerevisiae*. It has been reported that the conversion of PCA to catechol by PCA decarboxylase is a limiting step. For this reason we expressed the *Kp*AroY.B and *Kp*AroY.Ciso genes in either single or multiple genomic integrations to create a small

library of CCM production strains. In addition, Tkl1 was overexpressed in order to improve the precursor supply.

To establish a naringenin producing strain we integrated the full pathway containing the phenylalanine ammonium lyase from *Arabidopsis thaliana* (*At*PAL-2), the fusion of cinnamate 4-hydroxylase from *Arabidopsis thaliana* and NADPH-cytochrome P450 reductase from *Arabidopsis thaliana* (*At*C4H:L5:*At*ATR2), the 4-coumarate-CoA ligase 2 from *Arabidopsis thaliana* (*At*4CL-2), the naringenin-chalcone synthase from *Hypericum androsaemum* (*Ha*CHS), and the chalcone isomerase from *Petunia hybrida* (*Ph*CHI) to make strain EVR1 (table 1). Strains EVR2 and EVR3 contained one and two additional integrations of bottleneck enzymes (*At*PAL-2 and *Ha*CHS for EVR2; *At*PAL-2, *Ha*CHS, and *At*C4H:L5:*At*ATR2 for EVR3)(table 1).

Mutagenesis of BenM and library preparation. For optimization of the CCM inducibility of BenM, purified products from three consecutive rounds of error-prone PCR (epPCR) of the effector-binding domain (EBD, residues 90-304) of BenM, were co-transformed into yeast together with linearized centromeric plasmid according to Eckert-Boulet *et al.* (Eckert-Boulet, N., Pedersen, M. L., Krogh, B. O. & Lisby, M. Optimization of ordered plasmid assembly by gap repair in Saccharomyces cerevisiae. Yeast 29, 323–34 (2012)), to allow for *in vivo* gap repair and library reconstitution of wild-type BenM DNA binding domain (DBD) fused to EBD variants expressed from REV1p. For epPCR we used the GeneMorph II kit according to manufacturer's description (Agilent Technologies). Transformed yeast contained a chromosomal integration of the 209bp_CYC1p_*BenO*_T1 promoter controlling the expressing of GFP at EasyClone site 4 on chromosome XII (Mikkelsen, M. D. et al. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metab. Eng. 14, 104–11 (2012)), to allow for FACS-based screening of improved CCM-inducible BenM variants.

Metabolite quantification using HPLC and UPLC-MS. The CCM production strains were cultivated in 24-deep well plate with air-penetrable lids (EnzyScreen) to test for the production of CCM. Colonies from the individual strain were inoculated in 1mL synthetic drop-out medium (Sigma-Aldrich), selecting for *URA*, *HIS* and *LEU* markers, and grown at 30°C with 250 rpm agitation at 5 cm orbit cast for 24 h. 300µL of the overnight cultures were used to inoculate 3mL mineral medium (pH 4.5) in 24-deep well plate and incubated for 24-72 h at the same conditions as above. Experiments were performed in triplicates. The culture broth was centrifuged 3,500 × rpm and the supernatant analyzed for CCM concentration using HPLC. For this purpose, samples were analyzed for 45 min using Aminex HPX-87H ion exclusion column with a 1 mM H_2SO_4 flow of 0.6 mL/min. The temperature of the column was 60°C. Refractive index and UV detectors (Dionex) were used for detection of CCM (250 nm).

CCM concentrations were quantified by comparison with the spectrum of the standards. For the naringenin production strains 300µl culture broth was extracted with 300 µl MeOH in a 10-minute incubation (300 × rpm, 5cm amplitude, 30°C) in a 96 square deep-well microtiter plate (Greiner Masterblock, 96 Well, 2ml, P, V-bottom) and subsequently clarified by centrifugation at 4000 × g for 5 min. Clarified broth extract was subsequently diluted four times with 50% MeOH and 2µl was injected on a Acquity UPLC system (Waters) coupled to a Acquity TQ mass detector (Waters). Separation of the compounds was achieved on a Acquity UPLC® BEH C18 column (Waters, 1.7 µm, 2.1 mm x 50 mm), kept at 55°C . Mobile phases A and B were water containing 0.1 % formic acid and acetonitrile containing 0.1% formic acid, respectively. A flow of 0.6 ml/min was used. The gradient profile was as follow: 0.3 min constant at 10% B, a linear gradient from 10% B to 25% B in 3.7 min, a second linear gradient from 25% B to 100% B in 1 min, a wash for 1 min at 100% B and back to the initial condition of 10 % B for 0.6 min. The mass analyzer was equipped with an electrospray (ESI) source and operated in negative mode. Capillary voltage was 3.0kV; the source was kept at 150°C and the desolvation temperature was 350°C; desolvation and cone gas flow were 500 L/h and 50 L/h respectively. $[M-H]^-$ ions of naringenin (271 m/z) was tracked in SIR mode. Naringenin was quantified using a quadratic calibration curve with authentic standards ranging from 0.039 mg/l to 20 mg/l using QuanLynx software (Waters).

Transport assays. Overnight grown CEN.PK113-5A cells were diluted (OD_{600}=0.1) into SC-HIS-LEU medium with or without 1.4 mM (200mg/L) CCM. Media samples were taken at both 0 h and 24 h, while samples for measuring cellular lysates (10^8 cells) were harvested at 24 h. For quantification of CCM by LC-MS, cultures were harvested by centrifugation. For extracellular CCM quantification the supernatant was centrifuged twice and filtered (0.2µm) before analysis. For intracellular CCM quantification harvested pellets were washed twice in ice-cold isotonic saline solution (0.9% NaCl) and centrifuged at 5,000 x g before cells were extracted in an aqueous 0.1% formic acid solution and sonicated for 15 min. Following this, samples were centrifuged at 13,000 x g and supernatants filtered (0.2µm) prior to analysis. LC-MS data was collected on EVOQ EliteTriple Quadrupole Mass Spectrometer system coupled with an Advance UHPLC (Bruker). Samples were held at 4°C during the analysis. A 1µL sample was injected onto a ACQUITY HSS T3 C18 UHPLC column (Waters), with a 1.8µm particle size and 2.1 x 100 mm dimensioning. The column was held at a temperature of 30°C. The solvent system used was 0.1% formic acid (mobile phase A) and acetonitrile with 0.1% formic (mobile phase B). The flow rate was 0.400 ml/min with an initial solvent composition of 100% mobile phase A held until 0.50 min, then changed until it reached %A = 5.0 and %B = 95.0 at 1.00 min. This was held until 1.79 min when the solvent was returned to the initial conditions and the column was re-equilibrated until 4.00 min. The eluent was sprayed into the heated ESI probe of the MS which was held at 250°C and a voltage of 2500 V. Sheath/nebulizer/cone gas flow rate of 50/50/20 units and cone temp was 350°C. Two

transitions were chosen in negative Multiple Reaction Monitoring (MRM) mode for quantification of CCM: m/z 141.70-96.80 (quantification transition) and m/z 141.70-53.1(confirmation transition). Triplicate measurements were made for all samples.

Fluorescence Activated Cell Sorting

A two-step method was used to sort for BenM variants that specifically induce in the presence of CCM. Cells (10x library size, approx. 400,000 cells) were inoculated in mineral medium without inducer and incubated for 24 h at 30°C, diluted into PBS, and then GFP intensity of individual cells was measured using a BD Biosciences Aria (Becton Dickinson) with a blue laser (488 nm) by applying tight gates on the FSC and SSC channels. Only cells displaying auto-fluorescence intensities were sorted in order to limit auto-activating BenM variants. Sorted cells were recovered in mineral medium, followed by subculturing (1:100) into mineral medium containing 1.4 mM (200 mg/L) *cis,cis*-muconic acid. The cells were grown for 24 h at 30°C, washed and subjected to a second round of FACS. Cells exhibiting high levels of GFP (top 1%) were sorted, recovered in mineral medium and plated for single colonies on SC-HIS-LEU media. Individual clones were subsequently validated using flow cytometry.

Flow cytometry measurements and data analysis. Cells grown for 24 h in control (mineral medium, pH 4.5) or inducing medium (mineral medium pH 4.5 + 1.4 mM CCM, 1.4 mM protocatechuic acid, 10 mM malonic acid, 0.2 mM naringenin, or 50 mM L-arginine) were diluted into PBS to arrest cell growth. Cells were then analyzed by flow cytometry using a Fortessa flow cytometer (Becton Dickinson) with a blue laser (488 nm), for validation of single strains. For each strain 10,000 single-cell events were recorded. Events were analyzed using FlowJo software (TreeStar Inc.). The fluorescence arithmetic mean of the gated cell population was calculated, and the fold-change determined by dividing the mean fluorescence of the induced (ON) state with the mean fluorescence of the control (OFF) state. For flow cytometry for CCM and naringenin producing cells we tested mean fluorescence intensities from 10,000 cells pr. strain following 72 and 48 h, respectively. The data represent the average of three (n=3) biological replicates and error bars correspond to the standard deviation between these measurements.

Transcriptome analysis. To study the impact of ligand-induced BenM on genome-wide gene expression, triplicate cultures of strains MeLS0138 and MeLS0284 were grown for 24 h at 30°C in 50 ml mineral medium pH 4.5 with 1.4 mM CCM. Following this, total RNA was extracted essentially as previously described (Kildegaard, K. R. et al. Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance. Metab. Eng. 26, 57–66 (2014)). Briefly, 15 ml samples of the six cultures were harvested into a pre-chilled 50 ml tube with crushed ice and then immediately centrifuged at 4°C, 4000 x rpm for 5 min.

Subsequently, the pellets were resuspended in 2 ml of RNAlater® Solution (Ambion, Life Technologies) and incubated on ice for 1 h. Next, cells were pelleted by centrifugation (12,000 x rpm for 10 s) and transferred to liquid nitrogen, and stored at −80°C until further analysis. Total RNA extraction was performed using RNeasy® Mini Kit (QIAGEN). For this purpose, samples were thawed on ice, and 600 µl of buffer RLT containing 1% (v/v) β-mercaptoethanol was added directly to the cells, before being transferred into a 2 ml extraction tube containing 500µl glass beads and disrupted using the PRECELLYS®24 (Bertin Technologies) for 2 x 20 s at 6500 x rpm. The cell mixture was pelleted and supernatant transferred to a new tube. Total RNA was purified according to the manufacturing's protocol, and genomic DNA removed using Turbo DNA-free™ Kit (Ambion). The quantity and quality of the RNA samples were measured using Qubit 2.0 Fluorometer using the Qubit RNA BR Assay (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer using the RNA 6000 Nano Kit (Agilent Technologies), respectively. For sequencing we used 3µg of total RNA as input for TruSeq® Stranded mRNA Sample Preparation kit prior to sequencing on the MiSeq System using MiSeq Reagent Kit v3 150 cycles at a 2 × 75 bp read length configuration (Illumina) obtaining 38 M reads.

Bioinformatic resources. Two-dimensional heatmap plots were generated using the plot3D package and the R GUI. For ribbon structure representation of CCM-binding domain of BenM ^{H110,F211V,Y286N} the UCSF Chimera software was used (Pettersen, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–12 (2004)). For RNA-seq data analysis, TopHat (2.0.13) and Cufflinks (2.2.1) suite was employed as previously described (Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–78 (2012)). Expression levels (Fragments Per Kilobase of exon per Million fragments mapped: FPKM) from three (n=3) biological replicates of the conditions tested are processed with cuffdiff to obtain fold change differences and to perform statistical testing. A q-value cutoff of <0.05 was used to identify genes that have significant differential expression. Additionally, a >2-fold cut-off selection criterion was applied. Reference genome and annotations for CEN.PK113-7D strain were retrieved from Saccharomyces Genome Database (SGD; http://www.yeastgenome.org/). Genes with FPKM=0 for any replicate were removed from consideration.

Database for RNA-seq data. RNA-seq data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-4836 (http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-4836).

Table 1. Strain names and genotypes of all strains generated in this study. Strain names for those shown in Fig. 1c are ordered according to basal activity (see also Supplementary Table 2)

Strain name	Yeast Integrative Plasmid (parent strain)	Genotype
CEN.PK113-7D	-	mat a URA3 HIS3 LEU2 TRP1
CEN.PK102-5B	-	mat a ura3 his3 leu2
Cen.PK113-5A	-	mat a his3 leu2 trp1
MeLS0081	pMeLS0045 + pCfB0262 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 SpHIS5
MeLS0153	pCfB0262 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 SpHIS5
MeLS0079	pMeLS0044 + pCfB0262 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 SpHIS5
MeLS0152	pCfB0262 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 SpHIS5
MeLS0131	pMeLS0077 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 491bp_CYC1p::yEGFP-SpHIS5
MeLS0132	pMeLS0078 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 272bp_CYC1p::yEGFP-SpHIS5
MeLS0133	pMeLS0079 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 249bp_CYC1p::yEGFP-SpHIS5
MeLS0134	pMeLS0080 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 209bp_CYC1p::yEGFP-SpHIS5
MeLS0135	pMeLS0019 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 491bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0136	pMeLS0081 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 272bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0137	pMeLS0082 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 249bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0138	pMeLS0025 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU 209bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0177	pMeLS0020 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 491bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0259	pMeLS0086 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 272bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0260	pMeLS0088 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 249bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0180	pMeLS0026 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 209bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0178	pMeLS0021 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 491bp_CYC1p_BenO_T1/T2::yEGFP-SpHIS5
MeLS0261	pMeLS0087 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 272bp_CYC1p_BenO_T1/T2::yEGFP-SpHIS5
MeLS0262	pMeLS0089 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 249bp_CYC1p_BenO_T1/T2::yEGFP-SpHIS5
MeLS0181	pMeLS0027 + pCfB257 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 209bp_CYC1p_BenO_T1/T2::yEGFP-SpHIS5
MeLS0164	pMeLS0077 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 491bp_CYC1p::yEGFP-SpHIS5
MeLS0165	pMeLS0078 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 272bp_CYC1p::yEGFP-SpHIS5
MeLS0166	pMeLS0079 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 249bp_CYC1p::yEGFP-SpHIS5
MeLS0167	pMeLS0080 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 209bp_CYC1p::yEGFP-SpHIS5
MeLS0156	pMeLS0077 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 491bp_CYC1p::yEGFP-SpHIS5
MeLS0157	pMeLS0078 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 272bp_CYC1p::yEGFP-SpHIS5
MeLS0158	pMeLS0079 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 249bp_CYC1p::yEGFP-SpHIS5
MeLS0159	pMeLS0080 + pMeLS0044	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2

	(Cen.PK113-5A)	209bp_CYC1p::yEGFP-SpHIS5
MeLS0160	pMeLS0077 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 491bp_CYC1p::yEGFP-SpHIS5
MeLS0161	pMeLS0078 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 272bp_CYC1p::yEGFP-SpHIS5
MeLS0162	pMeLS0079 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 249bp_CYC1p::yEGFP-SpHIS5
MeLS0163	pMeLS0080 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 209bp_CYC1p::yEGFP-SpHIS5
MeLS0139	pMeLS0077 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 491bp_CYC1p::yEGFP-SpHIS5
MeLS0140	pMeLS0078 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 272bp_CYC1p::yEGFP-SpHIS5
MeLS0141	pMeLS0079 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 249bp_CYC1p::yEGFP-SpHIS5
MeLS0142	pMeLS0080 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 209bp_CYC1p::yEGFP-SpHIS5
MeLS0172	pMeLS0019 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0147	pMeLS0081 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0173	pMeLS0082 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0150	pMeLS0025 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0168	pMeLS0019 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0148	pMeLS0081 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0169	pMeLS0082 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0025	pMeLS0025 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0170	pMeLS0019 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0149	pMeLS0081 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0171	pMeLS0082 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0151	pMeLS0025 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0143	pMeLS0019 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0144	pMeLS0081 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0145	pMeLS0082 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0049	pMeLS0025 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1::yEGFP-SpHIS5
MeLS0263	pMeLS0020 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 491bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0190	pMeLS0086 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 272bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0192	pMeLS0088 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 249bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0264	pMeLS0026 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 209bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0020	pMeLS0020 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 491bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0182	pMeLS0086 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 272bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0184	pMeLS0088 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 249bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0026	pMeLS0026 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 209bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0265	pMeLS0020 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 491bp_CYC1p_BenO_T2::yEGFP-SpHIS5

MeLS0194	pMeLS0086 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 272bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0196	pMeLS0088 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 249bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0266	pMeLS0026 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 209bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0044	pMeLS0020 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 491bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0186	pMeLS0086 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 272bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0188	pMeLS0088 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 249bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0050	pMeLS0026 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 209bp_CYC1p_BenO_T2::yEGFP-SpHIS5
MeLS0267	pMeLS0021 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0191	pMeLS0087 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0193	pMeLS0089 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0268	pMeLS0027 + pMeLS0046 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0021	pMeLS0021 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0183	pMeLS0087 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0185	pMeLS0089 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0027	pMeLS0027 + pMeLS0044 (Cen.PK113-5A)	mat a his3 leu2 trp1 TEF1p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0269	pMeLS0021 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0195	pMeLS0087 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0197	pMeLS0089 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 249bp_CYC1p_BenO_T/2::yEGFP-SpHIS5
MeLS0270	pMeLS0027 + pMeLS0053 (Cen.PK113-5A)	mat a his3 leu2 trp1 RNR2p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0045	pMeLS0021 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 491bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0187	pMeLS0087 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 272bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0189	pMeLS0089 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 249bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0051	pMeLS0027 + pMeLS0045 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM-KI.LEU2 209bp_CYC1p_BenO_T1/2::yEGFP-SpHIS5
MeLS0284	pMeLS0025 + pMeLS0123 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::BenM(H110R, F211V, Y286N)-KI.LEU2 209bp_CYC1p_BenO_T1::yEGFP-SpHIS5
ST2377 (CCM intermediate ; with ScTkl1)	pCfB1237 + pCfB1239 (CEN.PK102-5B)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2
ST3054 (CCM intermediate ; no ScTkl1)	pCfB1237 + pCfB2695 (CEN.PK102-5B)	mat a ura3 his3 leu2 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2
ST3034 (CCM-multiple; with ScTkl1)	pCfB1241 (ST2377)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 TDH3p::PaAroZ + TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n
ST3059 (CCM-single; with ScTkl1)	pCfB2696 (ST2377)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2 + TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3
ST3058 (CCM-multiple; no	pCfB1241 (ST3054)	mat a ura3 his3 leu2 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n

ScTkl1)		
ST3154 (CCM-single; no ScTkl1)	pCfB2696 (ST3054)	mat a ura3 his3 leu2 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2 + TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3
ST4240-1 (Reference strain + biosensor)	pCfB2553 + pCfB2764 (CEN.PK113-7D)	mat a URA3 HIS3 LEU2 TRP1 209bp_CYC1p_BenO_T1::yEGFP-HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4241-1 (CCM-single; no ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3154)	mat a ura3 his3 leu2 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA KlLEU2 + TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3 + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4242-8 (CCM-single; with ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3059)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2 + TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3 + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4243-1 (CCM-multiple; no ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3058)	mat a ura3 his3 leu2 TEF1p::KpAroY.D-SpHIS5 + TDH3p::PaAroZ TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4244-1 (CCM-multiple; with ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3034)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 TDH3p::PaAroZ + TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4244-2 (CCM-multiple; with ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3034)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 TDH3p::PaAroZ + TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4245-2 (CCM-multiple; with ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3034)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 TDH3p::PaAroZ + TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
ST4246-2 (CCM-multiple; with ScTkl1 + biosensor)	pCfB2553 + pCfB2764 (ST3034)	mat a ura3 his3 leu2 TDH3p::ScTkl1 TEF1p::KpAroY.D-SpHIS5 TDH3p::PaAroZ + TEF1p::CaCatA-KlLEU2 + (TDH3p:: KpAroY.B TEF1p::KpAroY.Ciso-KlURA3tag) x n + 209bp_CYC1p_BenO_T1::yEGFP HphMXsyn + REV1p::BenM(H110R, F211V, Y286N)-KanMXsyn
TISNO-64	pTS-27 + pTS-21 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::PcaQ-KI.LEU2 209bp_CYC1p_PcaO_T1::yEGFP-SpHIS5
TISNO-66	pTS-29 + pTS-23 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::ArgP-KI.LEU2 209bp_CYC1p_ArgO_T1::yEGFP-SpHIS5
TISNO-67	pTS-30 + pTS-24 + pTS-39 (Cen.PK113-5A)	mat a his3 leu2 trp1 REV1p::MdcR-KI.LEU2 209bp_CYC1p_MdcO_T1::yEGFP-SpHIS5 TEF1pr::SpMAE1
TISNO-71	pTS-33 + pTS-21 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::PcaQ-KI.LEU2 209bp_CYC1p_PcaO_T1::yEGFP-SpHIS5
TISNO-73	pTS-35 + pTS-23 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::ArgP-KI.LEU2 209bp_CYC1p_ArgO_T1::yEGFP-SpHIS5
TISNO-74	pTS-36 + pTS-24 + pTS-39 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::MdcR-KI.LEU2 209bp_CYC1p_MdcO_T1::yEGFP-SpHIS5 TEF1pr::SpMAE1
TISNO-79	pCfB257 + pTS-21 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 209bp_CYC1p_PcaO_T1::yEGFP-SpHIS5
TISNO-81	pCfB257 + pTS-23 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 209bp_CYC1p_ArgO_T1::yEGFP-SpHIS5
TISNO-82	pCfB257 + pTS-24 + pTS-39 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 209bp_CYC1p_MdcO_T1::yEGFP-SpHIS5 TEF1pr::SpMAE1
TISNO-83	pCfB257 + pCfB2226 + pTS-37 (Cen.PK113-5A)	mat a his3 leu2 trp1 KI.LEU2 Sp.HIS5syn 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn
TISNO-89	pTS-36 + pTS-23 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::MdcR-KI.LEU2 209bp_CYC1p_ArgO_T1::yEGFP-SpHIS5
TISNO-90	pTS-35 + pTS-24 + pTS-39 (Cen.PK113-5A)	mat a his3 leu2 trp1 TDH3p::ArgP-KI.LEU2 209bp_CYC1p_MdcO_T1::yEGFP-SpHIS5 TEF1pr::SpMAE1

TISNO-93	pTS-45 + pTS-37 (Cen.PK102-5B)	mat a ura3 his3 leu2 TDH3p::FdeR-KI.URA3syn 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn
TISNO-95	pMije0124 + pTS-37 (Cen.PK102-5B)	mat a ura3 his3 leu2 REV1p::FdeR-KI.LEU2 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn
EV0	-	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0
EV1	pROP280 + pROP266 + pROP273 (EV0)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS TEF1p::PhCHI PDC1p::At4Cl
EV2	pROP338 + pROP339 + pROP191 (EV1)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS TEF1p::PhCHI PDC1p::At4Cl + TDH3p::AtPAL2 PGK1p::HaCHS
EV3	pROP423 + pROP339 + pVAN968 (EV2)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS TEF1p::PhCHI PDC1p::At4Cl + TDH3p::AtPAL2 PGK1p::HaCHS + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS
EVR0 (ctrl)	pCfB2198 + pTS-49 (EV0)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn1 + [CEN/ARS/URA3/TDH3p::FdeR]
EVR1	pCfB2198 +pTS-49 (EV1)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS TEF1p::PhCHI PDC1p::At4Cl + 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn1 + [CEN/ARS/URA3/TDH3p::FdeR]
EVR2	pCfB2198 + pTS-49 (EV2)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS TEF1p::PhCHI PDC1p::At4Cl + TDH3p::AtPAL2 PGK1p::HaCHS + 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn1 + [CEN/ARS/URA3/TDH3p::FdeR]
EVR3	pCfB2198 + pTS-49 (EV3)	mat a ura3::LoxP-KanMX-LoxP pad1-fdc1::LoxP-NATMX-LoxP aro10Δ0 + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS TEF1p::PhCHI PDC1p::At4Cl + TDH3p::AtPAL2 PGK1p::HaCHS + TDH3p::AtPAL2 TEF2p::C4H::L5::ATR2 PGK1p::HaCHS + 209bp_CYC1p_FdeO_T1::yEGFP-hphMXsyn1 + [CEN/ARS/URA3/TDH3p::FdeR]

Table 2. Mean fluorescense intensities of GFP (24 h)

Strain ID	Design	Ctrl	CCM (1.4 mM)	sd ctrl	sd CCM	FC (+/- CCM)
MeLS0081	REV1p-BenM	159	175	3	7	1,105042017
MeLS0153	RNR2p-BenM	154	174	2	2	1,130151844
MeLS0079	TEF1p-BenM	173	185	4	12	1,071428571
MeLS0152	TDH3p-BenM	170	184	12	3	1,080392157
MeLS0136	272bp_CYC1p_BenO_T1:GFP	155	160	0	1	1,032258065
MeLS0162	RNR2p-BenM + 249bp_CYC1p:GFP	146	164	2	4	1,120728929
MeLS0163	RNR2p-BenM + 209bp_CYC1p:GFP	146	164	4	3	1,120728929
MeLS0138	209bp_CYC1p_BenO_T1:GFP	147	167	5	4	1,133484163
MeLS0189	REV1p-BenM + 249bp_CYC1p_BenO_T1/T2:GFP	148	181	3	8	1,220224719
MeLS0181	209bp_CYC1p_BenO_T1/T2:GFP	150	165	7	6	1,102222222
MeLS0134	209bp_CYC1p:GFP	151	171	4	4	1,129955947
MeLS0180	209bp_CYC1p_BenO_T2:GFP	152	165	3	5	1,087719298

MeLS0159	TEF1p-BenM + 209bp_CYC1p:GFP	156	177	6	7	1,132478632
MeLS0137	249bp_CYC1p_BenO_T1:GFP	156	180	1	1	1,151385928
MeLS0050	REV1p-BenM + 209bp_CYC1p_BenO_T2:GFP	156	182	6	4	1,166311301
MeLS0158	TEF1p-BenM + 249bp_CYC1p:GFP	157	175	1	2	1,116772824
MeLS0142	REV1p-BenM + 209bp_CYC1p:GFP	159	174	1	2	1,089958159
MeLS0188	REV1p-BenM + 249bp_CYC1p_BenO_T2:GFP	160	194	5	5	1,214583333
MeLS0195	RNR2p-BenM + 272bp_CYC1p_BenO_T1/T2:GFP	161	199	6	7	1,238095238
MeLS0197	RNR2p-BenM + 249bp_CYC1p_BenO_T1/T2:GFP	163	256	2	15	1,570552147
MeLS0186	REV1p-BenM + 272bp_CYC1p_BenO_T2:GFP	164	197	5	3	1,201219512
MeLS0140	REV1p-BenM + 272bp_CYC1p:GFP	166	183	5	3	1,104627767
MeLS0141	REV1p-BenM + 249bp_CYC1p:GFP	167	176	8	3	1,055888224
MeLS0266	RNR2p-BenM + 209bp_CYC1p_BenO_T2:GFP	167	179	5	3	1,071856287
MeLS0161	RNR2p-BenM + 272bp_CYC1p:GFP	171	180	11	8	1,048638132
MeLS0132	272bp_CYC1p:GFP	174	211	6	7	1,213051823
MeLS0157	TEF1p-BenM + 272bp_CYC1p:GFP	175	198	8	26	1,12952381
MeLS0167	TDH3p-BenM + 209bp_CYC1p:GFP	178	194	4	12	1,090056285
MeLS0166	TDH3p-BenM + 249bp_CYC1p:GFP	181	205	1	10	1,130514706
MeLS0182	TEF1p-BenM + 272bp_CYC1p_BenO_T2:GFP	183	211	5	11	1,155109489
MeLS0165	TDH3p-BenM + 272bp_CYC1p:GFP	187	216	2	21	1,156862745
MeLS0171	RNR2p-BenM + 249bp_CYC1p_BenO_T1:GFP	187	290	6	4	1,549019608
MeLS0051	REV1p-BenM + 209bp_CYC1p_BenO_T1/T2:GFP	196	307	5	3	1,568994889
MeLS0070	RNR2p-BenM + 209bp_CYC1p_BenO_T1/T2:GFP	200	356	3	14	1,775374376
MeLS0190	TDH3p-BenM + 272bp_CYC1p_BenO_T2:GFP	201	218	7	7	1,088039867
MeLS0259	272bp_CYC1p_BenO_T2:GFP	201	201	4	12	1,001658375
MeLS0262	249bp_CYC1p_BenO_T1/T2:GFP	203	203	7	5	1
MeLS0187	REV1p-BenM + 272bp_CYC1p_BenO_T1/T2:GFP	207	370	7	4	1,789049919
MeLS0145	REV1p-BenM + 249bp_CYC1p_BenO_T1:GFP	207	405	2	17	1,953376206
MeLS0133	249bp_CYC1p:GFP	210	219	11	20	1,042789223
MeLS0196	RNR2p-BenM + 249bp_CYC1p_BenO_T2:GFP	221	349	7	10	1,575301205
MeLS0151	RNR2p-BenM + 209bp_CYC1p_BenO_T1:GFP	230	859	6	23	3,734782609
MeLS0049	REV1p-BenM + 209bp_CYC1p_BenO_T1:GFP	231	880	37	6	3,80952381
MelS0194	RNR2p-BenM + 272bp_CYC1p_BenO_T2:GFP	242	258	9	13	1,064649243
MeLS0178	491bp_CYC1p_BenO_T1/T2:GFP	248	241	4	3	0,97311828
MeLS0183	TEF1p-BenM + 272bp_CYC1p_BenO_T1/T2:GFP	259	511	4	11	1,969151671
MeLS0184	TEF1p-BenM + 249bp_CYC1p_BenO_T2:GFP	287	475	2	8	1,653132251
MeLS0261	272bp_CYC1p_BenO_T1/T2:GFP	321	257	17	2	0,798755187
MeLS0144	REV1p-BenM + 272bp_CYC1p_BenO_T1:GFP	338	804	4	23	2,381046397
MeLS0260	249bp_CYC1p_BenO_T2:GFP	344	313	11	20	0,909002904

MeLS0264	TDH3p-BenM + 209bp_CYC1p_BenO_T2:GFP	353	527	7	5	1,495274102
MeLS0020	TEF1p-BenM + 491bp_CYC1p_BenO_T2:GFP	365	813	14	18	2,224452555
MeLS0191	TDH3p-BenM + 272bp_CYC1p_BenO_T1/T2:GFP	365	691	12	43	1,892335766
MeLS0149	RNR2p-BenM + 272bp_CYC1p_BenO_T1:GFP	369	854	21	15	2,313459801
MeLS0026	TEF1p-BenM + 209bp_CYC1p_BenO_T2:GFP	380	406	5	8	1,068481124
MeLS0193	TDH3p-BenM + 249bp_CYC1p_BenO_T1/T2:GFP	407	1630	10	67	4,004095004
MeLS0044	REV1p-BenM + 491bp_CYC1p_BenO_T2:GFP	438	667	18	7	1,523990861
MeLS0265	RNR2p-BenM + 491bp_CYC1p_BenO_T2:GFP	461	698	22	21	1,51300578
MeLS0027	TEF1p-BenM + 209bp_CYC1p_BenO_T1/T2:GFP	554	1879	60	20	3,393738712
MeLS0192	TDH3p-BenM + 249bp_CYC1p_BenO_T2:GFP	555	978	13	22	1,760504202
MeLS0263	TDH3p-BenM + 491bp_CYC1p_BenO_T2:GFP	631	1197	15	15	1,897517169
MeLS0177	491bp_CYC1p_BenO_T2:GFP	641	698	18	7	1,088403536
MeLS0268	TDH3p-BenM + 209bp_CYC1p_BenO_T1/T2:GFP	668	2172	14	96	3,250374065
MeLS0169	TEF1p-BenM + 249bp_CYC1p_BenO_T1:GFP	672	2913	12	21	4,337468983
MeLS0135	491bp_CYC1p_BenO_T1:GFP	754	801	12	11	1,061864781
MeLS0173	TDH3p-BenM + 249bp_CYC1p_BenO_T1:GFP	755	3554	6	78	4,706843267
MeLS0267	TDH3p-BenM + 491bp_CYC1p_BenO_T1/T2:GFP	840	2218	33	81	2,639825466
MeLS0148	TEF1p-BenM + 272bp_CYC1p_BenO_T1:GFP	1164	4032	39	145	3,463212139
MeLS0025	TEF1p-BenM + 209bp_CYC1p_BenO_T1:GFP	1185	4139	6	85	3,493528419
MeLS0150	TDH3p-BenM + 209bp_CYC1p_BenO_T1:GFP	1192	4868	16	47	4,082471345
MeLS0045	REV1p-BenM + 491bp_CYC1p_BenO_T1/T2:GFP	1197	1698	8	67	1,418151448
MeLS0269	RNR2p-BenM + 491bp_CYC1p_BenO_T1/T2:GFP	1291	2589	80	74	2,005422153
MeLS0147	TDH3p-BenM + 272bp_CYC1p_BenO_T1:GFP	1333	4984	28	65	3,738
MeLS0021	TEF1p-BenM + 491bp_CYC1p_BenO_T1/T2:GFP	1358	2967	16	142	2,18404908
MeLS0139	REV1p-BenM + 491bp_CYC1p:GFP	2270	2231	1813	1771	0,98281686
MeLS0160	RNR2p-BenM + 491bp_CYC1p:GFP	2625	2811	39	118	1,07059421
MeLS0185	TEF1p-BenM + 249bp_CYC1p_BenO_T1/T2:GFP	2632	9992	273	294	3,796833439
MeLS0131	491bp_CYC1p:GFP	2781	2899	30	116	1,042435867
MeLS0156	TEF1p-BenM + 491bp_CYC1p:GFP	3046	3016	44	67	0,990152095
MeLS0164	TDH3p-BenM + 491bp_CYC1p:GFP	4003	4509	90	105	1,126488467
MeLS0143	REV1p-BenM + 491bp_CYC1p_BenO_T1:GFP	5321	5789	181	229	1,0880842
MeLS0170	RNR2p-BenM + 491bp_CYC1p_BenO_T1:GFP	5616	6114	87	48	1,088734568
MeLS0172	TDH3p-BenM + 491bp_CYC1p_BenO_T1:GFP	6894	10605	24	353	1,53834252
MeLS0168	TEF1p-BenM + 491bp_CYC1p_BenO_T1:GFP	15429	20250	355	316	1,312441938

Table 3. List of plasmids

Plasmid name	Parent plasmid	Description	Reference/Source
pCfB258	-	pX-4-LoxP-SpHIS5	Jensen et al., 2014
pCfB322	-	pTY4-LoxP-KlURA3tag	Borodina et al., 2014
pCfB388	-	pXI-1-LoxP-KlLEU2	Jensen et al., 2014
pCfB390	-	pXI-3-LoxP-KlURA3	Jensen et al., 2014
pCfB2198	-	pXII-4-LoxP-HphMXsyn	Stovicek et al., 2015
pCfB2223	-	pX-3-LoxP-KanMXsyn	Stovicek et al., 2015
pCfB2226	-	pX-4-loxP-SpHIS5syn	Stovicek et al., 2015
pCfB1237	pCfB258	pX-4-loxP-SpHiS5-ScTkl1<-TDH3p-TEF1p->KpAroY.D	This study
pCfB1239	pCfB388	pXI-1-LoxP-KlLEU2-PaAroZ<- TDH3p-TEF1p->CaCatA	This study
pCfB1241	pCfB322	pTY4-LoxP-KlURA3tag-KpAroY.B<- TDH3p-TEF1p->KpAroY.Ciso	This study
pCfB2374	-	pXI-1-loxP-KlURA3syn	Stovicek et al., 2015
pCfB3039	-	pXII-2	Stovicek et al., 2015
pCfB2695	pCfB258	pX-4-LoxP-SpHiS5-TEF1p->KpAroY.D	This study
pCfB2696	pCfB390	pXI-3-KlURA3-KpAroY.B<- TDH3p-TEF1p->KpAroY.Ciso	This study
pCfB2553	pCfB2198	pXII-4-LoxP-HphMXsyn-209bp_CYC1p_BenO_T1->yEGFP	This study
pCfB2764	pCfB2223	pX-3-LoxP-KanMXsyn-REV1p->BenM(H110R, F211V, Y286N)	This study
pCfB257	-	pX-3-LoxP-KlLEU2	Jensen et al., 2014
pCfB262	-	pXII-4-LoxP-SpHIS5	Jensen et al., 2014
pRS416U	-	URA3, USER cassette	This study
pMeLS0045	pCfB257	pX-3-LoxP-KlLEU2-REV1p->BenM	This study
pMeLS0044	pCfB257	pX-3-LoxP-KlLEU2-TEF1p->BenM	This study
pMeLS0053	pCfB257	pX-3-LoxP-KlLEU2-RNR2p->BenM	This study
pMeLS0046	pCfB257	pX-3-LoxP-KlLEU2-TDH3p->BenM	This study
pMeLS0077	pCfB262	pXII-4-LoxP-SpHIS5-491bp_CYC1p->yeGFP	This study
pMeLS0078	pCfB262	pXII-4-LoxP-SpHIS5-272bp_CYC1p->yeGFP	This study
pMeLS0079	pCfB262	pXII-4-LoxP-SpHIS5-249bp_CYC1p->yeGFP	This study
pMeLS0080	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p->yeGFP	This study
pMeLS0019	pCfB262	pXII-4-LoxP-SpHIS5-491bp_CYC1p_BenO_T1->yeGFP	This study
pMeLS0081	pCfB262	pXII-4-LoxP-SpHIS5-272bp_CYC1p_BenO_T1->yeGFP	This study
pMeLS0082	pCfB262	pXII-4-LoxP-SpHIS5-249bp_CYC1p_BenO_T1->yeGFP	This study
pMeLS0025	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p_BenO_T1->yeGFP	This study
pMeLS0020	pCfB262	pXII-4-LoxP-SpHIS5-491bp_CYC1p_BenO_T2->yeGFP	This study
pMeLS0086	pCfB262	pXII-4-LoxP-SpHIS5-272bp_CYC1p_BenO_T2->yeGFP	This study
pMeLS008	pCfB262	pXII-4-LoxP-SpHIS5-249bp_CYC1p_BenO_T2-	This study

8		>yeGFP	
pMeLS0026	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p_BenO_T2->yeGFP	This study
pMeLS0021	pCfB262	pXII-4-LoxP-SpHIS5-491bp_CYC1p_BenO_T1/2->yeGFP	This study
pMeLS0087	pCfB262	pXII-4-LoxP-SpHIS5-272bp_CYC1p_BenO_T1/2->yeGFP	This study
pMeLS0089	pCfB262	pXII-4-LoxP-SpHIS5-249bp_CYC1p_BenO_T1/2->yeGFP	This study
pMeLS0027	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p_BenO_T1/2->yeGFP	This study
pMeLS0123	pCfB257	pX-3-LoxP-KlLEU2-REV1p->BenM(H110R, F211V, Y286N)	This study
pMije0124	pCfB257	pX-3-LoxP-KlLEU2-REV1p->FdeR	This study
pTS-21	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p_PcaO_T1->yeGFP	This study
pTS-23	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p_ArgO_T1->yeGFP	This study
pTS-24	pCfB262	pXII-4-LoxP-SpHIS5-209bp_CYC1p_MdcO_T1->yeGFP	This study
pTS-27	pCfB257	pX-3-LoxP-KlLEU2-REV1p->PcaQ	This study
pTS-29	pCfB257	pX-3-LoxP-KlLEU2-REV1p->ArgP	This study
pTS-30	pCfB257	pX-3-LoxP-KlLEU2-REV1p->MdcR	This study
pTS-33	pCfB257	pX-3-LoxP-KlLEU2-TDH3p->PcaQ	This study
pTS-35	pCfB257	pX-3-LoxP-KlLEU2-TDH3p->ArgP	This study
pTS-36	pCfB257	pX-3-LoxP-KlLEU2-TDH3p->MdcR	This study
pTS-37	pCfB2198	pXII-4-LoxP-HphMX-209bp_CYC1p_FdeO_T1->yeGFP	This study
pTS-38	pCfB2374	pXI-1-LoxP-Kl.URA3syn-TDH3p->FdeR	This study
pTS-39	pCfB3039	pXII-2-TEF1p->Sp.MAE1	This study
pTS-49	pRS416U	pURA3-TDH3p->FdeR	This study
pROP280	-	pX-DR-KlLEU2-AtPAL2<-TDH3p-TEF2p->AtC4H::L5::AtATR2	This study
pROP266	-	HaCHS<-PGK1p-TEF1p->PhCHI	This study
pROP273	-	pX-PDC1p->At4CL2	This study
pROP338	-	pXI-DR-KlLEU2-AtPAL2<-TDH3p	This study
pROP339	-	HaCHS<-PGK1p	This study
pROP191	-	pXI	This study
pROP423	-	pXVI-DR-KlLEU2-AtPAL2<-TDH3p-TEF2p->AtC4H::L5::AtATR2	This study
pVAN968	-	pXVI	This study

Table 4. List of synthetic genes. For synthetic reporter promoters based on the 209bp_CYC1p scaffold promoter, LTTR operator sites are marked in bold and start codon for yeGFP is marked in green.

```
LOCUS       CaCatA                  912 bp

       1  ATGTCCCAAG CTTTCACCGA ATCTGTTAAG ACTTCTTTGG GTCCAAATGC TACTCCAAGA
      61  GCTAAAAAGT TGATTGCCTC TTTGGTTCAA CACGTTCATG ATTTCGCTAG AGAAAACCAT
     121  TTGACTACCG AAGATTGGTT GTGGGGTGTT GATTTCATTA ACAGAATTGG TCAAATGTCC
     181  GACTCCAGAA GAAACGAAGG TATTTTGGTT TGCGATATCA TCGGTTTGGA AACCTTGGTT
     241  GATGCTTTGA CTAACGAATC CGAACAATCT AACCATACCT CCTCTGCTAT TTTGGGTCCT
```

```
     301  TTTTACTTGC CAGATTCTCC AGTTTATCCA AACGGTGGTT CTATCGTTCA AAAGGCTATT
     361  CCAACTGATG TTAAGTGCTT CGTTAGAGGT AAGGTTACTG ATACTGAAGG TAAACCATTG
     421  GGTGGTGCTC AATTGGAAGT TTGGCAATGT AATTCTGCTG GTTTCTACTC TCAACAAGCT
     481  GATCATGATG GTCCAGAATT CAATTTGAGA GGTACTTTCA TTACCGACGA CGAAGGTAAT
     541  TACTCCTTCG AATGTTTAAG ACCAACCTCC TATCCAATTC CATACGATGG TCCTGCTGGT
     601  GATTTGTTGA AAATCATGGA TAGACATCCA AACAGACCAT CCCATATTCA TTGGAGAGTT
     661  TCTCATCCAG GTTACCATAC TTTGATCACC CAAATCTATG ATGCTGAATG TCCATACACC
     721  AACAACGATT CTGTTTACGC TGTTAAGGAT GACATCATCG TTCACTTCGA AAAGGTTGAT
     781  AACAAGGATA AGGATTTGGT CGGTAAGGTC GAATACAAGT TGGATTACGA TATTTCCTTG
     841  GCCACCGAAT CCTCTATTCA AGAAGCTAGA GCTGCTGCTA AAGCTAGACA AGATGCTGAA
     901  ATCAAGTTGT AA
//
LOCUS       KpAroY.B                 594 bp

       1  ATGAAGTTGA TCATCGGTAT GACTGGTGCT ACAGGTGCTC CATTGGGTGT TGCTTTGTTG
      61  CAAGCTTTGA GAGATATGCC AGAAGTTGAA ACCCATTTGG TTATGTCTAA ATGGGCTAAG
     121  ACCACCATTG AATTGGAAAC TCCATGGACT GCTAGAGAAG TTGCTGCTTT GGCTGATTTT
     181  TCTCATTCTC CAGCTGATCA AGCTGCTACT ATTTCTTCTG GTTCTTTCAG AACTGATGGT
     241  ATGATCGTTA TTCCATGCTC TATGAAAACC TTGGCTGGTA TTAGAGCTGG TTATGCTGAA
     301  GGTTTGGTTG GTAGAGCTGC TGATGTTGTT TTGAAAGAAG GTAGAAAGTT GGTCTTGGTC
     361  CCAAGAGAAA TGCCATTGTC TACTATCCAT TTGGAAAACA TGTTGGCCTT GTCTAGAATG
     421  GGTGTAGCTA TGGTTCCACC AATGCCAGCT TATTACAATC ATCCAGAAAC CGTTGATGAC
     481  ATCACCAACC ATATAGTTAC CAGAGTTTTG GACCAATTCG GTTTGGATTA TCACAAAGCT
     541  AGAAGATGGA ACGGTTTGAG AACTGCTGAA CAATTCGCTC AAGAAATTGA ATGA
//
LOCUS       KpAroY.Ciso              1509 bp

       1  ATGACCGCCC CAATCCAAGA TTTGAGAGAT GCTATTGCTT TGTTACAACA ACACGACAAT
      61  CAATACTTGG AAACCGATCA TCCAGTTGAT CCAAATGCTG AATTGGCTGG TGTTTACAGA
     121  CATATTGGTG CTGGTGGTAC TGTAAAAAGA CCAACTAGAA TTGGTCCAGC CATGATGTTC
     181  AACAACATTA AGGGTTATCC ACACTCCAGA ATCTTGGTTG GTATGCATGC TTCTAGACAA
     241  AGAGCAGCTT TGTTGTTGGG TTGTGAAGCT TCTCAATTGG CTTTGGAAGT TGGTAAAGCT
     301  GTTAAGAAAC CAGTTGCTCC AGTTGTTGTT CCAGCTTCTT CTGCTCCATG TCAAGAACAA
     361  ATTTTCTTGG CTGATGATCC AGACTTCGAT TTGAGAACTT TGTTGCCAGC TCATACCAAC
     421  ACTCCAATTG ATGCTGGTCC ATTTTTTTGT TTGGGTTTGG CTTTAGCTTC TGATCCTGTT
     481  GATGCTTCTT TGACCGATGT TACCATTCAT AGATTGTGCG TTCAAGGTAG AGATGAATTG
     541  TCTATGTTTT TGGCTGCCGG TAGACATATC GAAGTTTTTA GACAAAAAGC TGAAGCTGCT
     601  GGTAAGCCAT TGCCAATTAC TATTAACATG GGTTTAGATC CAGCCATCTA CATTGGTGCT
     661  TGTTTTGAAG CTCCAACTAC TCCATTTGGT TACAACGAAT TGGGTGTTGC TGGTGCTTTG
     721  AGACAAAGAC CAGTTGAATT GGTTCAAGGT GTTTCTGTTC CAGAAAAGGC TATTGCTAGA
     781  GCCGAAATAG TTATCGAAGG TGAATTATTG CCAGGTGTCA GAGTTAGAGA AGATCAACAT
     841  ACAAATTCCG GTCATGCTAT GCCAGAATTT CCAGGTTATT GTGGTGGTGC TAATCCATCT
     901  TTGCCAGTTA TTAAGGTTAA GGCCGTTACC ATGAGAAACA ACGCTATTTT ACAAACTTTG
     961  GTCGGTCCAG GTGAAGAACA TACAACTTTG GCTGGTTTGC CAACCGAAGC TTCTATTTGG
    1021  AATGCTGTTG AAGCTGCAAT TCCAGGTTTC TTGCAAAATG TTTATGCTCA TACAGCTGGT
    1081  GGTGGTAAGT TCTTGGGTAT ATTGCAAGTC AAGAAAAGAC AACCAGCTGA CGAAGGTAGA
    1141  CAAGGTCAAG CTGCTTTATT AGCTTTGGCT ACTTACTCCG AATTGAAGAA TATCATCTTG
    1201  GTCGATGAAG ATGTTGATAT CTTCGATTCC GATGATATTT TGTGGGCTAT GACTACTAGA
    1261  ATGCAAGGTG ATGTTTCCAT TACTACCATT CCAGGTATTA GAGGTCACCA ATTAGATCCA
    1321  TCTCAAACCC CAGAATACTC CCCATCAATT AGAGGTAATG GTATCTCCTG TAAGACCATT
    1381  TTCGATTGCA CTGTTCCATG GGCTTTGAAG TCTCATTTTG AAAGAGCACC ATTTGCTGAC
    1441  GTTGATCCTA GACCTTTTGC TCCAGAATAT TTCGCTAGAT TGGAAAAGAA TCAAGGTTCC
    1501  GCTAAGTAA
//
LOCUS       KpAroY.D                 237 bp

       1  ATGATCTGTC CAAGATGCGC CGACGAAAAA ATTGAAGTTA TGGCTACTTC TCCAGTTAAG
      61  GGTGTTTGGA CTGTTTATCA ATGTCAACAC TGCTTGTACA CTTGGAGAGA TACTGAACCA
     121  TTGAGAAGAA CCTCTAGAGA ACATTACCCT GAAGCTTTCA GAATGACCCA AAAGGATATT
     181  GATGAAGCTC CACAAGTTCC TCATGTTCCA CCATTATTGC CAGAAGATAA GAGATAA
//
LOCUS       PaAroZ                   1104 bp

       1  ATGCCATCCA AGTTGGCCAT TACCTCTATG TCTTTGGGTA GATGTTATGC CGGTCATTCT
      61  TTCACTACTA AGTTGGATAT GGCTAGAAAG TACGGTTACC AAGGTTTGGA ATTATTCCAT
     121  GAAGATTTGG CTGATGTCGC CTATAGATTG TCTGGTGAAA CTCCATCTCC ATGTGGTCCA
     181  TCACCAGCTG CTCAATTGTC TGCTGCTAGA CAAATTTTGA GAATGTGCCA AGTCAGAAAC
     241  ATCGAAATCG TTTGCTTGCA ACCATTCTCT CAATACGATG GTTTGTTGGA TAGAGAAGAA
     301  CACGAAAGAA GATTGGAACA ATTGGAATTC TGGATCGAAT TGGCCCATGA ATTGGATACC
     361  GATATTATTC AAATTCCAGC CAACTTCTTG CCAGCCGAAG AAGTTACTGA AGATATCTCT
     421  TTGATTGTCT CCGACTTGCA AGAAGTAGCT GATATGGGTT TACAAGCTAA CCCACCAATT
     481  AGATTCGTTT ACGAAGCTTT GTGTTGGTCC ACTAGAGTTG ATACTTGGGA AAGATCTTGG
     541  GAAGTTGTTC AAAGAGTTAA CAGACCAAAC TTCGGTGTTT GTTTGGACAC TTTTAACATT
     601  GCCGGTAGAG TTTATGCTGA TCCAACTGTT GCTTCTGGTA GAACTCCAAA TGCTGAAGAA
```

```
     661  GCTATCAGAA AGTCCATTGC CAGATTGGTT GAAAGAGTTG ACGTTTCCAA GGTTTTCTAC
     721  GTTCAAGTTG TTGATGCCGA AAAGTTGAAG AAACCATTGG TTCCAGGTCA CAGATTCTAT
     781  GATCCAGAAC AACCAGCTAG AATGTCTTGG TCTAGAAACT GCAGATTATT CTACGGTGAA
     841  AAGGATAGAG GTGCTTACTT GCCAGTAAAA GAAATTGCTT GGGCTTTTTT CAACGGTTTG
     901  GGTTTTGAAG GTTGGGTTTC CTTAGAATTA TTCAACAGAA GAATGTCCGA TACCGGTTTT
     961  GGTGTTCCAG AAGAATTAGC TAGAAGAGGT GCTGTTTCTT GGGCTAAATT GGTTAGAGAT
    1021  ATGAAGATCA CCGTTGACTC TCCAACTCAA CAACAAGCTA CACAACAACC TATCAGAATG
    1081  TTGTCTTTGT CAGCTGCTTT GTGA
//
LOCUS       BenM                       915 bp

       1  ATGGAATTGA GACACTTGAG ATACTTCGTT GCCGTTGTTG AAGAACAATC TTTTACAAAG
      61  GCTGCCGACA AGTTGTGTAT TGCTCAACCA CCATTATCCA GACAAATCCA AAACTTGGAA
     121  GAAGAATTGG GTATCCAATT ATTGGAAAGA GGTTCCAGAC CAGTTAAGAC TACTCCAGAA
     181  GGTCATTTCT TTTACCAATA CGCCATCAAG TTGTTGTCCA ACGTTGATCA AATGGTCAGT
     241  ATGACCAAGA GAATTGCCTC TGTTGAAAAG ACCATTAGAA TCGGTTTTGT TGGTTCCTTG
     301  TTGTTCGGTT TGTTGCCAAG AATTATCCAC TTGTACAGAC AAGCTCATCC AAACTTGAGA
     361  ATCGAATTAT ACGAAATGGG TACTAAGGCT CAAACCGAAG CTTTGAAAGA AGGTAGAATT
     421  GACGCTGGTT TTGGTAGATT GAAGATTTCT GATCCAGCCA TCAAGAGAAC CTTGTTGAGA
     481  AACGAAAGAT TGATGGTTGC TGTTCATGCT TCCCATCCAT TGAATCAAAT GAAGGATAAG
     541  GGTGTTCACT TGAACGATTT GATCGACGAA AAGATCTTGT TGTACCCATC TTCTCCAAAG
     601  CCAAACTTCT CTACTCATGT TATGAACATC TTCTCTGACC ATGGTTTGGA ACCTACCAAG
     661  ATTAACGAAG TTAGAGAAGT CCAATTGGCC TTGGGTTTGG TTGCTGCTGG TGAAGGTATT
     721  TCATTGGTTC CAGCTTCTAC CCAATCCATT CAATTATTCA ACTTGTCCTA CGTCCCATTA
     781  TTAGATCCAG ATGCTATTAC CCCAATCTAC ATTGCTGTTA GAAACATGGA AGAATCCACC
     841  TACATCTACT CATTATACGA AACCATCAGA CAAATCTACG CCTACGAAGG TTTTACTGAA
     901  CCACCAAATT GGTAA
//
LOCUS       FdeR                       930 bp

       1  ATGCGTTTCA ACAAGCTCGA CCTCAATCTT CTGGTCGCCC TGGATGCACT GCTCACGGAG
      61  ATGAGCATCA GCCGCGCCGC CGAAAAGATC CATCTGAGCC AGTCGGCCAT GAGCAATGCC
     121  CTGGCGCGGC TGCGCGAGTA TTTCGATGAT GAATTGCTGA TCCAGGTGGG CCGGCGCATG
     181  GAGCCCACGC CGCGCGCCGA GGTGCTCAAG GATGCGGTGC ATGATGTGCT GCGGCGTATC
     241  GATGGCTCCA TCGCGGCGCT GCCGGCCTTC GTGCCGGCCG AGTCCACGCG CGAGTTTCGC
     301  ATCTCGGTTT CGGACTTTAC GCTCTCCGTC CTCATCCCCC GGGTGCTGGC GCGCGCGCAC
     361  GCCGAGGGCA AGCACATCCG CTTTGCCCTG ATGCCGCAGG TGCAAGACCC GACCCGCTCG
     421  CTGGATCGGG CCGAGGTGGA CCTGCTGGTC TTGCCGCAGG AATTCTGCAC GCCCGATCAT
     481  CCTGCCGAAG AGGTCTTCCG CGAACGGCAT GTCTGCGTGG TCTGGCGCGA CAGTGCGCTG
     541  GCGCAAGGCG AGCTGACGCT GGAACGCTAC ATGGCCTCAG GCCATGTGGT GATGGTGCCG
     601  CCTGGGGCCA ATGCGTCGTC GGTGGAGGCG TGGATGGCCA GGAAGCTGGG CTTTGCGCGC
     661  CGGGTGGAAG TGACCAGCTT CAGCTTCGCT TCTGCGCTGG CGCTGGTACA GGGGACGGAC
     721  CGCATCGCCA CGGTGCATGC CCGGCTGGCG CAGCTGCTGG CTCCGCAATG GCCGGTGGTG
     781  ATCAAGGAGA GTCCGCTGTC GCTGGGCGAG ATGCGGCAGA TGATGCAGTG GCATCGCTAC
     841  CGCAGCAATG ATCCTGGCAT CCAGTGGCTG CGTCGGGTGT TTCTGGAGAG TGCGCAGGAG
     901  ATGGATGCGG CGCTGCCAGG CATCTGCTGA
//
LOCUS       PcaQ             942 bp

       1  ATGATTGATG CACGTGTGAA ATTTAGACAT TTGCAAACTT TTGTAGAAGT TGCTAGACAA
      61  AAGAGTGTTG TAAAAGCAGC CGAATTATTA CATGTAACAC AGCCAGCAGT GACTAAGACC
     121  ATAAGGGAAT TGGAAGAGGT ATTAGGTGTC GCCGTGTTTG AAAGAGAAGG TCGTGGTATC
     181  AAAATAACAA GATATGGGGA AGTTTTTTTG AGACATGCAG GAGCTGCCCT TACGGCTCTT
     241  CGTCAAGGTC TAGACAGCGT ATCTCAAGAA AGAAGTGGCG AAGGTCCACC AATCAGGGTA
     301  GGCGCCTTAC CTACAGTATC AACTAGAATC ATGCCAAGAG CTATTGCACT TTTTCTGAAG
     361  GAAAAAACGG GTGCAAGAAT TAAAATAGTC ACAGGCGAAA ATGCGGTATT GCTTGAACAA
     421  TTGAGAATCG GCGACCTAGA CTTGGTTGTG GGAAGGCTTG CCGCCCCGGA TAAAATGACT
     481  GGGTTTTCTT TCGAGCACCT ATACAGTGAG CAAGTTGTGT TTGCAGTAAG GGCAGGCCAT
     541  CCCCTGATCT CCGGTAGGCA ATCCTTGTTT GCTCATCTTT CCGACTACCC TGTTCTAATG
     601  CCAACAAGGG CCAGCATAAT TAGGCCATTC GTCGAGCACT TTTTGATAGC TAATGGCATC
     661  GCTGGTTTGC CAAACCAGAT AGAAACCGTC TCCGATTCAT TTGGTAGAGC TTTTGTACGT
     721  TCTTCCGACG CTATTTGGAT TATATCCGCT GGTGTAGTAG CTACTGATAT TGCCGATGGT
     781  GTTTTGGCAG CTCTACCAGT AGACACTTCA GAAACCCGTG GCCCTGTTGG CTTGACTATG
     841  AGAACCGATG CAATACCATC TTTGCCTCTT TCAATCTTAA TGCAAACTTT AAGAGAAGTG
     901  GCCGGTACCG CAATGGCAGC TGAAGCCAAA AGAACAGCAT AA
//
LOCUS       ArgP             894 bp    DNA

       1  ATGAAACGTC CTGATTATAG AACTCTGCAA GCCTTAGATG CTGTAATTAG AGAACGTGGC
      61  TTCGAGAGAG CGGCTCAGAA GTTGTGTATT ACTCAATCCG CCGTGAGCCA GAGAATAAAG
     121  CAGCTAGAAA ATATGTTTGG CCAACCATTA CTGGTACGTA CTGTTCCTCC TAGGCCGACG
     181  GAACAAGGTC AGAAGCTTTT GGCCTTGTTG AGACAAGTGG AGTTGCTAGA AGAGGAATGG
     241  TTGGGAGACG AGCAGACAGG TTCAACACCA CTTTTATTGA GTCTGGCCGT AAATGCGGAT
     301  AGCCTAGCTA CTTGGTTGCT ACCGGCTCTA GCTCCTGTCT TGGCTGACAG TCCCATAAGA
```

```
     361  TTAAACTTAC AAGTCGAAGA TGAAACGAGA ACGCAAGAAA GACTTAGGAG AGGAGAGGTC
     421  GTGGGGGCTG TATCAATTCA ACATCAGGCA TTGCCCAGTT GTTTAGTCGA CAAGTTGGGT
     481  GCGCTAGATT ACCTTTTCGT GTCATCCAAA CCTTTCGCCG AGAAGTATTT TCCTAATGGC
     541  GTTACCCGTT CCGCTTTGCT TAAAGCCCCA GTCGTAGCAT TCGACCATCT AGATGACATG
     601  CACCAAGCCT TTTTACAACA AAATTTCGAT TTACCACCAG GCTCCGTTCC ATGCCATATC
     661  GTGAACTCTT CCGAAGCCTT CGTACAACTA GCTCGTCAAG GTACTACTTG CTGTATGATT
     721  CCACATCTAC AAATAGAAAA AGAATTGGCC TCCGGAGAAT TGATAGACCT GACACCTGGC
     781  CTATTTCAAA GGAGAATGCT GTATTGGCAT AGGTTTGCAC CAGAGTCAAG AATGATGAGA
     841  AAGGTGACTG ATGCATTGCT TGATTATGGC CATAAGGTGT TAAGACAAGA TTGA
//
LOCUS       MdcR            927 bp

       1  ATGAAGGACG ACATCAATCA AGAAATTACC TTCAGGAAGT TATCTGTTTT CATGATGTTT
      61  ATGGCCAAAG GCAATATCGC CAGAACTGCT GAAGCAATGA AGTTATCATC TGTGTCAGTT
     121  CACAGAGCGC TGCATACACT AGAAGAAGGT GTGGGATGTC CCCTGTTCGT CCACAAAGGT
     181  AGAAATCTAC TACCTCTACA GGCAGCATGG ACTCTATTAG AATATTGCCA AGATGTAATT
     241  TCATTAATGA ATAGAGGACT AGAAGCCACT AGAAAAGTGG CAGGTGTTGG TCAAGGAAGA
     301  TTGAGAATCG GTACACTTTA CTCCTTAACA CTAGAAACCG TACCAAGGAT AATAATGGGC
     361  ATGAAGTTAA GACGTCCAGA ACTTGAGCTA GACTTGACAA TGGGTTCAAA TCAAATGTTA
     421  TTAGATATGC TAGAAGATGA TGCCTTAGAT GCAATATTGA TAGCTACCAA CGAAGGCGAA
     481  TTCAACAATA CTGCCTTTGA TGTTGTTCCT TTGTTTGAGG ATGACATATT TCTTGCAGCA
     541  CCTGCAACTG AACGTCTTGA CGCCTCAAGA TTGGCTGACC TGAGAGATTA CGCTGATAGA
     601  AAGTTTGTTT CCTTAGCGGA AGGATTTGCT ACCTATGCTG GTTTTCGTGA AGCTTTCCAT
     661  ATAGCTGGCT TTGAACCAGA GATAGTTACC AGAGTTAATG ACATATTCAG TATGATATCT
     721  CTTGTTCAGG CTGGTGTTGG GTTTGCTCTT TTGCCAGGAA GAATGAAGAA AGTTTATGAA
     781  AAGGACGTTC AATTGCTTAA GTTAGCCGAA CCTTACCAAA TGAGACAGCT GATTAGTATC
     841  GTATATTCCC ATCACAGGGA ACGTGACGCT GATTTGTTGG CATTAGCGGC TGAAGGTAGG
     901  ATGTATGCTC GTTCTATTAA CAGGTAA
//
LOCUS       209bp_CYC1p_BenO_T1::yeGFP          1014 bp

       1  CCAGGCAACT TTAGTGCTGA CACATAATAC TCCATAGGTA TTTTATTATA CAAATAATGT
      61  GTTTGAACTT ATTAAAACAT TCTTTTAAGG TATAAACAAC AGGCATATAT ATATGTGTGC
     121  GACGACACAT GATCATATGG CATGCATGTG CTCTGTATGT ATATAAAACT CTTGTTTTCT
     181  TCTTTTCTCT AAATATTCTT TCCTTATACA TTAGGACCTT TGCAGCATAA ATTACTATAC
     241  TTCTATAGAC ACACAAACAC AAATACACAC ACTAAATTAA TAATCTGTCA TAAAACAATG
     301  TCTAAAGGTG AAGAATTATT CACTGGTGTT GTCCCAATTT TGGTTGAATT AGATGGTGAT
     361  GTTAATGGTC ACAAATTTTC TGTCTCCGGT GAAGGTGAAG GTGATGCTAC TTACGGTAAA
     421  TTGACCTTAA AATTTATTTG TACTACTGGT AAATTGCCAG TTCCATGGCC AACCTTAGTC
     481  ACTACTTTCG GTTATGGTGT TCAATGTTTT GCGAGATACC CAGATCATAT GAAACAACAT
     541  GACTTTTTCA AGTCTGCCAT GCCAGAAGGT TATGTTCAAG AAAGAACTAT TTTTTCAAA
     601  GATGACGGTA ACTACAAGAC CAGAGCTGAA GTCAAGTTTG AAGGTGATAC CTTAGTTAAT
     661  AGAATCGAAT TAAAAGGTAT TGATTTTAAA GAAGATGGTA ACATTTTAGG TCACAAATTG
     721  GAATACAACT ATAACTCTCA CAATGTTTAC ATCATGGCTG ACAAACAAAA GAATGGTATC
     781  AAAGTTAACT TCAAAATTAG ACACAACATT GAAGATGGTT CTGTTCAATT AGCTGACCAT
     841  TATCAACAAA ATACTCCAAT TGGTGATGGT CCAGTCTTGT TACCAGACAA CCATTACTTA
     901  TCCACTCAAT CTGCCTTATC CAAAGATCCA AACGAAAAGA GAGACCACAT GGTCTTGTTA
     961  GAATTTGTTA CTGCTGCTGG TATTACCCAT GGTATGGATG AATTGTACAA ATAA
//
LOCUS       209bp_CYC1p_FdeO_T1::yeGFP          1014 bp

       1  CCAGGCAACT TTAGTGCTGA CACATAAGCT TGATATTGAT CAAATGGATT GTTTTGATTC
      61  ATGATATGGA CGGCATCAAT ACATTGACCA CCCCATCCGC AGGCATATAT ATATGTGTGC
     121  GACGACACAT GATCATATGG CATGCATGTG CTCTGTATGT ATATAAAACT CTTGTTTTCT
     181  TCTTTTCTCT AAATATTCTT TCCTTATACA TTAGGACCTT TGCAGCATAA ATTACTATAC
     241  TTCTATAGAC ACACAAACAC AAATACACAC ACTAAATTAA TAATCTGTCA TAAAACAATG
     301  TCTAAAGGTG AAGAATTATT CACTGGTGTT GTCCCAATTT TGGTTGAATT AGATGGTGAT
     361  GTTAATGGTC ACAAATTTTC TGTCTCCGGT GAAGGTGAAG GTGATGCTAC TTACGGTAAA
     421  TTGACCTTAA AATTTATTTG TACTACTGGT AAATTGCCAG TTCCATGGCC AACCTTAGTC
     481  ACTACTTTCG GTTATGGTGT TCAATGTTTT GCGAGATACC CAGATCATAT GAAACAACAT
     541  GACTTTTTCA AGTCTGCCAT GCCAGAAGGT TATGTTCAAG AAAGAACTAT TTTTTCAAA
     601  GATGACGGTA ACTACAAGAC CAGAGCTGAA GTCAAGTTTG AAGGTGATAC CTTAGTTAAT
     661  AGAATCGAAT TAAAAGGTAT TGATTTTAAA GAAGATGGTA ACATTTTAGG TCACAAATTG
     721  GAATACAACT ATAACTCTCA CAATGTTTAC ATCATGGCTG ACAAACAAAA GAATGGTATC
     781  AAAGTTAACT TCAAAATTAG ACACAACATT GAAGATGGTT CTGTTCAATT AGCTGACCAT
     841  TATCAACAAA ATACTCCAAT TGGTGATGGT CCAGTCTTGT TACCAGACAA CCATTACTTA
     901  TCCACTCAAT CTGCCTTATC CAAAGATCCA AACGAAAAGA GAGACCACAT GGTCTTGTTA
     961  GAATTTGTTA CTGCTGCTGG TATTACCCAT GGTATGGATG AATTGTACAA ATAA
//
LOCUS       209bp_CYC1p_PcaO_T1::yeGFP          1021 bp

       1  CCAGGCAACT TTAGTGCTGA CACATAGATC GTATAACCTC CTGGTTAAGG GAAAGCCACG
      61  AAATATCATT TTACCTAACC GGATGAAACA TCCAAATCTG ACGACGCAGG CATATATATA
     121  TGTGTGCGAC GACACATGAT CATATGGCAT GCATGTGCTC TGTATGTATA TAAAACTCTT
```

```
     181  GTTTTCTTCT TTTCTCTAAA TATTCTTTCC TTATACATTA GGACCTTTGC AGCATAAATT
     241  ACTATACTTC TATAGACACA CAAACACAAA TACACACACT AAATTAATAA TCTGTCATAA
     301  AACAATGTCT AAAGGTGAAG AATTATTCAC TGGTGTTGTC CCAATTTTGG TTGAATTAGA
     361  TGGTGATGTT AATGGTCACA AATTTTCTGT CTCCGGTGAA GGTGAAGGTG ATGCTACTTA
     421  CGGTAAATTG ACCTTAAAAT TTATTTGTAC TACTGGTAAA TTGCCAGTTC CATGGCCAAC
     481  CTTAGTCACT ACTTTCGGTT ATGGTGTTCA ATGTTTTGCG AGATACCCAG ATCATATGAA
     541  ACAACATGAC TTTTTCAAGT CTGCCATGCC AGAAGGTTAT GTTCAAGAAA GAACTATTTT
     601  TTTCAAAGAT GACGGTAACT ACAAGACCAG AGCTGAAGTC AAGTTTGAAG GTGATACCTT
     661  AGTTAATAGA ATCGAATTAA AAGGTATTGA TTTTAAAGAA GATGGTAACA TTTTAGGTCA
     721  CAAATTGGAA TACAACTATA ACTCTCACAA TGTTTACATC ATGGCTGACA AACAAAAGAA
     781  TGGTATCAAA GTTAACTTCA AAATTAGACA CAACATTGAA GATGGTTCTG TTCAATTAGC
     841  TGACCATTAT CAACAAAATA CTCCAATTGG TGATGGTCCA GTCTTGTTAC CAGACAACCA
     901  TTACTTATCC ACTCAATCTG CCTTATCCAA AGATCCAAAC GAAAAGAGAG ACCACATGGT
     961  CTTGTTAGAA TTTGTTACTG CTGCTGGTAT TACCCATGGT ATGGATGAAT TGTACAAATA
    1021  A
//
LOCUS       209bp_CYC1p_Arg0_T1::yeGFP          1028 bp

       1  CCAGGCAACT TTAGTGCTGA CACATATCTG GCCTCTCTCT TATTAGTTTT TCTGATTGCC
      61  AATTAATATT ATCAATTTCC GCTAATAACA ATCCCGCGAT ATAGTCTCTG CATCAGGCAT
     121  ATATATATGT GTGCGACGAC ACATGATCAT ATGGCATGCA TGTGCTCTGT ATGTATATAA
     181  AACTCTTGTT TTCTTCTTTT CTCTAAATAT TCTTTCCTTA TACATTAGGA CCTTTGCAGC
     241  ATAAATTACT ATACTTCTAT AGACACACAA ACACAAATAC ACACACTAAA TTAATAATCT
     301  GTCATAAAAC AATGTCTAAA GGTGAAGAAT TATTCACTGG TGTTGTCCCA ATTTTGGTTG
     361  AATTAGATGG TGATGTTAAT GGTCACAAAT TTTCTGTCTC CGGTGAAGGT GAAGGTGATG
     421  CTACTTACGG TAAATTGACC TTAAAATTTA TTTGTACTAC TGGTAAATTG CCAGTTCCAT
     481  GGCCAACCTT AGTCACTACT TTCGGTTATG GTGTTCAATG TTTTGCGAGA TACCCAGATC
     541  ATATGAAACA ACATGACTTT TTCAAGTCTG CCATGCCAGA AGGTTATGTT CAAGAAAGAA
     601  CTATTTTTTT CAAAGATGAC GGTAACTACA AGACCAGAGC TGAAGTCAAG TTTGAAGGTG
     661  ATACCTTAGT TAATAGAATC GAATTAAAAG GTATTGATTT TAAAGAAGAT GGTAACATTT
     721  TAGGTCACAA ATTGGAATAC AACTATAACT CTCACAATGT TTACATCATG GCTGACAAAC
     781  AAAAGAATGG TATCAAAGTT AACTTCAAAA TTAGACACAA CATTGAAGAT GGTTCTGTTC
     841  AATTAGCTGA CCATTATCAA CAAAATACTC CAATTGGTGA TGGTCCAGTC TTGTTACCAG
     901  ACAACCATTA CTTATCCACT CAATCTGCCT TATCCAAAGA TCCAAACGAA AAGAGAGACC
     961  ACATGGTCTT GTTAGAATTT GTTACTGCTG CTGGTATTAC CCATGGTATG GATGAATTGT
    1021  ACAAATAA
//
LOCUS       209bp_CYC1p_MdcO_T1::yeGFP          1030 bp

       1  CCAGGCAACT TTAGTGCTGA CACATAATCG TTACTCTGAT GCTAACGATC GGCCACCGCG
      61  CTTAATTGAT GCTCATAGCC TCGCGTCGCA CACTAATCTC CACCAGGACA AACAACAGGC
     121  ATATATATAT GTGTGCGACG ACACATGATC ATATGGCATG CATGTGCTCT GTATGTATAT
     181  AAAACTCTTG TTTTCTTCTT TTCTCTAAAT ATTCTTTCCT TATACATTAG GACCTTTGCA
     241  GCATAAATTA CTATACTTCT ATAGACACAC AAACACAAAT ACACACACTA AATTAATAAT
     301  CTGTCATAAA ACAATGTCTA AAGGTGAAGA ATTATTCACT GGTGTTGTCC CAATTTTGGT
     361  TGAATTAGAT GGTGATGTTA ATGGTCACAA ATTTTCTGTC TCCGGTGAAG GTGAAGGTGA
     421  TGCTACTTAC GGTAAATTGA CCTTAAAATT TATTTGTACT ACTGGTAAAT TGCCAGTTCC
     481  ATGGCCAACC TTAGTCACTA CTTTCGGTTA TGGTGTTCAA TGTTTTGCGA GATACCCAGA
     541  TCATATGAAA CAACATGACT TTTTCAAGTC TGCCATGCCA GAAGGTTATG TTCAAGAAAG
     601  AACTATTTTT TTCAAAGATG ACGGTAACTA CAAGACCAGA GCTGAAGTCA AGTTTGAAGG
     661  TGATACCTTA GTTAATAGAA TCGAATTAAA AGGTATTGAT TTTAAAGAAG ATGGTAACAT
     721  TTTAGGTCAC AAATTGGAAT ACAACTATAA CTCTCACAAT GTTTACATCA TGGCTGACAA
     781  ACAAAAGAAT GGTATCAAAG TTAACTTCAA AATTAGACAC AACATTGAAG ATGGTTCTGT
     841  TCAATTAGCT GACCATTATC AACAAAATAC TCCAATTGGT GATGGTCCAG TCTTGTTACC
     901  AGACAACCAT TACTTATCCA CTCAATCTGC CTTATCCAAA GATCCAAACG AAAAGAGAGA
     961  CCACATGGTC TTGTTAGAAT TTGTTACTGC TGCTGGTATT ACCCATGGTA TGGATGAATT
    1021  GTACAAATAA
//
LOCUS       SpMAE1              1317 bp

       1  ATGGGTGAAC TCAAGGAAAT CTTGAAACAG AGGTATCATG AGTTGCTTGA CTGGAATGTC
      61  AAAGCCCCTC ATGTCCCTCT CAGTCAACGA CTGAAGCATT TTACATGGTC TTGGTTTGCA
     121  TGTACTATGG CAACTGGTGG TGTTGGTTTG ATTATTGGTT CTTTCCCCTT TCGATTTTAT
     181  GGTCTTAATA CAATTGGCAA AATTGTTTAT ATTCTTCAAA TCTTTTTGTT TTCTCTCTTT
     241  GGATCATGCA TGCTTTTTCG CTTTATTAAA TATCCTTCAA CTATCAAGGA TTCCTGGAAC
     301  CATCATTTGG AAAAGCTTTT CATTGCTACT TGTCTTCTTT CAATATCCAC GTTCATCGAC
     361  ATGCTTGCCA TATACGCCTA TCCTGATACC GGCGAGTGGA TGGTGTGGGT CATTCGAATC
     421  CTTTATTACA TTTACGTTGC AGTATCCTTT ATATACTGCG TAATGGCTTT TTTTACAATT
     481  TTCAACAACC ATGTATATAC CATTGAAACC GCATCTCCTG CTTGGATTCT TCCTATTTTC
     541  CCTCCTATGA TTTGTGGTGT CATTGCTGGC GCCGTCAATT CTACACAACC CGCTCATCAA
     601  TTAAAAAATA TGGTTATCTT TGGTATCCTC TTTCAAGGAC TTGGTTTTTG GGTTTATCTT
     661  TTACTGTTTG CCGTCAATGT CTTACGGTTT TTTACTGTAG GCCTGGCAAA ACCCCAAGAT
     721  CGACCTGGTA TGTTTATGTT TGTCGGTCCA CCAGCTTTCT CAGGTTTGGC CTTAATTAAT
     781  ATTGCGCGTG GTGCTATGGG CAGTCGCCCT TATATTTTTG TTGGCGCCAA CTCATCCGAG
```

```
     841  TATCTTGGTT TTGTTTCTAC CTTTATGGCT ATTTTTATTT GGGGTCTTGC TGCTTGGTGT
     901  TACTGTCTCG CCATGGTTAG CTTTTTAGCG GGCTTTTTCA CTCGAGCCCC TCTCAAGTTT
     961  GCTTGTGGAT GGTTTGCATT CATTTTCCCC AACGTGGGTT TTGTTAATTG TACCATTGAG
    1021  ATAGGTAAAA TGATAGATTC CAAAGCTTTC CAAATGTTTG GACATATCAT TGGGGTCATT
    1081  CTTTGTATTC AGTGGATCCT CCTAATGTAT TTAATGGTCC GTGCGTTTCT CGTCAATGAT
    1141  CTTTGCTATC CTGGCAAAGA CGAAGATGCC CATCCTCCAC CAAAACCAAA TACAGGTGTC
    1201  CTTAACCCTA CCTTCCCACC TGAAAAAGCA CCTGCATCTT TGGAAAAAGT CGATACACAT
    1261  GTCACATCTA CTGGTGGTGA ATCGGATCCT CCTAGTAGTG AACATGAAAG CGTTTAA
//
LOCUS       AtPAL-2              2154 bp

       1  ATGGACCAAA TTGAAGCAAT GCTATGCGGT GGTGGTGAAA AGACCAAGGT GGCCGTAACG
      61  ACAAAAACTC TTGCAGATCC TTTGAATTGG GGTCTGGCAG CTGACCAGAT GAAAGGTAGC
     121  CATCTGGATG AAGTTAAGAA GATGGTTGAG GAATACAGAA GACCAGTCGT AAATCTAGGC
     181  GGCGAGACAT TGACGATAGG ACAGGTAGCT GCTATTTCGA CCGTTGGCGG TTCAGTGAAG
     241  GTAGAACTTG CAGAAACAAG TAGAGCCGGA GTTAAGGCTT CATCAGATTG GGTCATGGAA
     301  AGTATGAACA AGGGCACAGA TTCCTATGGC GTTACCACAG GCTTTGGTGC TACCTCTCAT
     361  AGAAGAACTA AAAATGGCAC TGCTTTGCAA ACAGAACTGA TCAGATTCCT TAACGCCGGT
     421  ATTTTCGGTA ATACAAAGGA AACTTGCCAT ACATTACCCC AATCGGCAAC AAGAGCTGCT
     481  ATGCTTGTTA GGGTGAACAC TTTGTTGCAA GGTTACTCTG GAATAAGGTT TGAAATTCTT
     541  GAGGCCATCA CTTCACTATT GAACCACAAC ATTTCTCCTT CGTTGCCCTT AAGAGGAACA
     601  ATAACTGCCA GCGGTGATTT GGTTCCCCTT TCATATATCG CAGGCTTATT AACGGGAAGA
     661  CCTAATTCAA AGGCCACTGG TCCAGACGGA GAATCCTTAA CCGCTAAGGA AGCATTTGAG
     721  AAAGCTGGTA TTTCAACTGG TTTCTTTGAT TTgCAACCCA AGGAAGGTTT AGCCCTGGTG
     781  AATGGCACCG CTGTCGGCAG CGGTATGGCA TCCATGGTGT TGTTTGAAGC TAACGTACAA
     841  GCAGTTTTGG CCGAAGTTTT GTCCGCAATT TTTGCCGAAG TCATGAGTGG AAAACCTGAG
     901  TTTACTGATC ACTTGACCCA CAGGTTAAAA CATCACCCAG GACAAATTGA AGCAGCAGCT
     961  ATCATGGAGC ACATTTTGGA CGGCTCTAGC TACATGAAGT TAGCCCAGAA GGTTCATGAA
    1021  ATGGACCCTT TGCAAAAACC CAAACAAGAT AGATATGCTT TAAGGACATC CCCACAATGG
    1081  CTTGGCCCTC AAATTGAAGT AATTAGACAA GCTACAAAGT CTATAGAAAG AGAGATCAAC
    1141  TCTGTTAACG ATAATCCACT TATTGATGTG TCGAGGAATA AGGCAATACA TGGAGGCAAT
    1201  TTCCAGGGTA CACCCATAGG AGTCAGTATG GATAATACCA GGCTTGCCAT AGCCGCAATT
    1261  GGCAAATTAA TGTTTGCCCA ATTTTCTGAA TTGGTCAATG ACTTCTACAA TAACGGTTTG
    1321  CCTTCGAATC TGACCGCATC TTCTAACCCT AGTCTTGATT ATGGTTTCAA AGGTGCTGAG
    1381  ATAGCAATGG CAAGCTATTG TTCAGAGCTG CAATATCTAG CCAACCCAGT AACCTCTCAT
    1441  GTACAATCAG CCGAACAACA CAATCAGGAT GTTAATTCTT TGGGCCTGAT TTCATCAAGA
    1501  AAAACAAGCG AGGCCGTTGA TATCCTTAAA TTAATGTCCA CAACATTTTT AGTGGGTATA
    1561  TGCCAGGCCG TAGATTTgAG ACACTTGGAA GAGAATTTGA GACAGACAGT GAAAAATACC
    1621  GTATCACAGG TTGCAAAAAA GGTTCTAACT ACAGGTATCA ATGGTGAATT GCACCCATCA
    1681  AGATTCTGTG AAAAAGATTT ATTAAAAGTT GTAGATAGAG AACAAGTATT TACTTACGTT
    1741  GACGATCCAT GTAGCGCTAC TTATCCATTG ATGCAGAGAT TGAGACAAGT TATTGTAGAT
    1801  CACGCTTTAT CCAATGGTGA AACTGAGAAA AATGCCGTTA CTTCAATATT CCAAAAGATA
    1861  GGTGCCTTTG AAGAAGAACT GAAGGCAGTT TTACCAAAGG AAGTCGAAGC TGCTAGAGCC
    1921  GCATACGGAA ATGGTACTGC CCCTATACCA AATAGAATCA AAGAGTGTAG GTCGTACCCT
    1981  TTGTACAGAT TCGTTAGAGA AGAGTTGGGA ACCAAATTAC TAACTGGTGA AAAAGTCGTT
    2041  AGCCCAGGTG AAGAATTTGA CAAGGTATTC ACAGCTATGT GCGAGGGAAA GTTGATAGAT
    2101  CCACTTATGG ATTGCTTGAA AGAGTGGAAT GGTGCACCTA TTCCAATCTG CTAA
//
LOCUS       AtC4H::L5::AtATR2      3702 bp

       1 ATGGATTTGT TATTGCTGGA AAAGTCACTT ATTGCTGTAT TTGTGGCAGT TATTCTAGCC
      61 ACGGTTATTT CTAAATTAAG AGGTAAGAAA CTAAAACTAC CTCCTGGTCC CATCCCCATA
     121 CCAATTTTTG GTAATTGGTT GCAAGTGGGC GATGATTTGA ATCACAGAAA TTTGGTAGAC
     181 TATGCTAAGA AGTTCGGTGA CCTTTTCTTG CTTAGAATGG GTCAAAGGAA TTTGGTAGTG
     241 GTTAGCTCAC CTGATTTGAC TAAGGAGGTC TTATTAACGC AAGGCGTTGA GTTTGGCTCC
     301 AGAACTAGAA ATGTTGTGTT TGATATTTTC ACTGGTAAAG GTCAAGATAT GGTTTTTACA
     361 GTTTACGGTG AGCACTGGAG AAAAATGAGA AGAATCATGA CCGTACCATT CTTTACTAAC
     421 AAGGTTGTTC AACAAAATAG AGAAGGTTGG GAGTTTGAGG CAGCTTCCGT AGTGGAAGAC
     481 GTAAAGAAAA ATCCAGATTC GGCCACAAAG GGTATAGTAC TAAGAAAAAG ACTACAATTG
     541 ATGATGTACA ACAATATGTT CAGAATTATG TTTGACAGAA GATTTGAAAG TGAAGATGAC
     601 CCTTTGTTCC TGAGACTTAA GGCTTTGAAT GGTGAAAGAT CGAGATTGGC TCAAAGTTTC
     661 GAATATAATT ACGGTGACTT TATTCCAATC TTAAGACCAT TTTTGAGAGG CTATTTGAAA
     721 ATTTGCCAAG ACGTCAAGGA TAGGAGGATC GCTCTTTTCA AGAAGTACTT TGTGGACGAG
     781 AGAAAGCAAA TAGCTTCTTC CAAGCCCACA GGTTCGGAAG GTTTAAAATG TGCAATTGAT
     841 CATATTTTAG AAGCTGAACA AAAAGGTGAA ATTAACGAAG ATAATGTTTT GTACATTGTA
     901 GAAAATATCA ATGTGGCTGC AATAGAAACA ACCTTATGGT CAATAGAATG GGGTATTGCT
     961 GAATTGGTGA ATCACCCAGA AATACAATCT AAACTGAGAA ACGAGCTAGA TACCGTTTTA
    1021 GGTCCAGGTG TCCAAGTTAC AGAACCTGAT TTGCATAAGT TACCCTACTT GCAAGCTGTG
    1081 GTTAAAGAAA CCTTGAGATT GAGAATGGCT ATTCCTCTTC TAGTTCCTCA TATGAACCTA
    1141 CATGATGCTA AACTGGCCGG TTATGATATT CCAGCAGAAA GTAAGATTTT AGTAAATGCA
    1201 TGGTGGTTGG CCAACAATCC AAACAGTTGG AAAAAGCCTG AAGAATTCAG ACCTGAAAGA
    1261 TTCTTCGAAG AGGAATCTCA TGTTGAAGCC AACGGAAATG ACTTCAGATA TGTACCTTTT
    1321 GGCGTTGGCA GAAGATCGTG TCCAGGAATA ATACTAGCCT TACCAATATT GGGTATCACA
    1381 ATTGGTAGGA TGGTTCAAAA TTTTGAGTTG CTACCACCAC CCGGACAATC GAAAGTCGAT
```

```
     1441 ACTTCAGAGA AAGGAGGACA ATTCTCATTG CATATTTTGA ATCATTCCAT TATAGTCATG
     1501 AAACCCAGAA ATTGTAGCGC TGAAGCTGCA GCAAAAGAAG CTGCAGCTAA AGAAGCTGCA
     1561 GCAAAAGCTT CCAGTAGCTC TTCCTCCTCA ACCTCGATGA TCGACTTAAT GGCTGCTATT
     1621 ATAAAAGGAG AACCAGTTAT AGTTAGTGAC CCTGCTAACG CAAGCGCTTA CGAATCCGTT
     1681 GCAGCCGAGT TGTCAAGTAT GCTTATAGAA AATAGACAGT TTGCTATGAT TGTAACGACC
     1741 AGCATCGCCG TTTTAATTGG TTGCATCGTG ATGTTGGTGT GGAGGAGGAG CGGTTCGGGC
     1801 AATTCAAAGA GGGTTGAACC ACTAAAGCCA TTAGTTATCA AACCTAGAGA AGAGGAAATT
     1861 GACGATGGAA GGAAGAAAGT CACTATATTC TTCGGCACCC AAACAGGTAC AGCTGAAGGT
     1921 TTTGCTAAGG CTCTAGGAGA AGAAGCAAAA GCTAGATATG AAAAGACGAG ATTCAAAATT
     1981 GTCGATCTGG ATGACTATGC CGCCGATGAT GACGAATACG AAGAAAAATT GAAGAAAGAA
     2041 GATGTCGCAT TTTTCTTCCT TGCCACCTAC GGCGACGGTG AACCAACAGA TAATGCCGCA
     2101 AGGTTTTACA AGTGGTTTAC TGAAGGTAAT GACAGAGGAG AATGGCTGAA GAATTTGAAA
     2161 TATGGTGTGT TCGGCCTTGG TAACAGACAG TACGAGCATT TTAATAAGGT CGCTAAGGTT
     2221 GTAGATGATA TACTTGTTGA ACAAGGTGCT CAAAGGTTAG TGCAGGTGGG CTTGGGTGAC
     2281 GATGATCAAT GTATTGAAGA TGACTTTACT GCTTGGAGAG AAGCCTTGTG GCCTGAATTA
     2341 GATACTATCC TTAGAGAAGA AGGTGACACT GCTGTTGCTA CCCCCTACAC TGCAGCAGTC
     2401 CTAGAATATA GAGTCTCAAT CCATGATTCA GAAGACGCCA AATTCAATGA TATTAACATG
     2461 GCCAACGGTA ACGGTTACAC CGTTTTTGAC GCACAACATC CATACAAAGC TAATGTTGCT
     2521 GTTAAAAGGG AACTTCACAC CCCAGAAAGT GACAGGTCAT GTATACATTT GGAATTTGAT
     2581 ATCGCTGGTA GTGGTTTGAC TTACGAAACA GGTGACCATG TCGGAGTACT TTGCGATAAT
     2641 TTGTCAGAAA CTGTTGATGA AGCTTTGAGG TTATTGGATA TGTCACCAGA TACTTACTTC
     2701 TCATTGCATG CAGAAAAAGA AGACGGAACT CCAATATCAA GCTCGCTTCC CCCTCCATTC
     2761 CCTCCCTGTA ACTTAAGAAC AGCCCTAACT AGATATGCTT GTTTACTGTC TTCTCCAAAG
     2821 AAAAGTGCTT TGGTTGCATT GGCAGCCCAC GCATCCGATC CTACCGAAGC TGAGAGATTA
     2881 AAGCATTTGG CTTCACCAGC CGGTAAAGAT GAATACAGTA AGTGGGTAGT GGAGAGCCAA
     2941 AGATCGCTTT TAGAAGTGAT GGCTGAGTTT CCAAGTGCTA AACCTCCTCT GGGTGTATTT
     3001 TTCGCTGGTG TGGCCCCAAG ATTGCAGCCT AGATTTTATT CCATATCCTC ATCTCCAAAA
     3061 ATTGCCGAAA CCAGAATTCA CGTGACATGT GCTCTGGTCT ACGAAAAGAT GCCAACAGGT
     3121 AGGATTCACA AGGGTGTCTG TTCTACCTGG ATGAAAAATG CTGTACCCTA TGAAAAATCC
     3181 GAAAATTGTT CTAGTGCACC AATTTTCGTA AGACAATCTA ATTTCAAGTT ACCAAGCGAT
     3241 TCTAAAGTAC CCATTATTAT GATCGGTCCA GGTACTGGTT TGGCCCCATT CAGAGGCTTC
     3301 TTGCAAGAAA GATTGGCTTT AGTGGAGAGT GGAGTTGAAT TGGGTCCTTC AGTTTTATTC
     3361 TTTGGTTGTA GAAACAGAAG AATGGACTTT ATCTACGAAG AAGAATTGCA GAGATTTGTT
     3421 GAAAGTGGTG CATTGGCCGA ATTGAGTGTT GCATTCAGCA GGGAAGGTCC AACCAAAGAA
     3481 TACGTTCAAC ACAAGATGAT GGACAAGGCT TCTGATATCT GGAATATGAT TTCCCAAGGT
     3541 GCTTATTTGT ATGTTTGTGG TGACGCTAAA GGAATGGCTA GAGATGTTCA TAGATCACTG
     3601 CATACAATCG CACAAGAACA AGGTAGCATG GATTCAACAA AAGCAGAGGG CTTTGTAAAG
     3661 AATCTTCAGA CAAGCGGTAG ATATCTGAGA GATGTATGGT AA
//
LOCUS       At4CL-2          1671 bp

        1  ATGACGACAC AAGATGTGAT AGTCAATGAT CAGAATGATC AGAAACAGTG TAGTAATGAC
       61  GTCATTTTCC GATCGAGATT GCCTGATATA TACATCCCTA ACCACCTCCC ACTCCACGAC
      121  TACATCTTCG AAAATATCTC AGAGTTCGCC GCTAAGCCAT GCTTGATCAA CGGTCCCACC
      181  GGCGAAGTAT ACACCTACGC CGATGTCCAC GTAACATCTC GGAAACTCGC CGCCGGTCTT
      241  CATAACCTCG GCGTGAAGCA ACACGACGTT GTAATGATCC TCCTCCCGAA CTCTCCTGAA
      301  GTAGTCCTCA CTTTCCTTGC CGCCTCCTTC ATCGGCGCAA TCACCACCTC CGCGAACCCG
      361  TTCTTCACTC CGGCGGAGAT TTCTAAACAA GCCAAAGCCT CCGCTGCGAA ACTCATCGTC
      421  ACTCAATCCC GTTACGTCGA TAAAATCAAG AACCTCCAAA ACGACGGCGT TTTGATCGTC
      481  ACCACCGACT CCGACGCCAT CCCCGAAAAC TGCCTCCGTT TCTCCGAGTT AACTCAGTCC
      541  GAAGAACCAC GAGTGGACTC AATACCGGAG AAGATTTCGC CAGAAGACGT CGTGGCGCTT
      601  CCTTTCTCAT CCGGCACGAC GGGTCTCCCC AAAGGAGTGA TGCTAACACA CAAAGGTCTA
      661  GTCACGAGCG TGGCGCAGCA AGTCGACGGC GAGAATCCGA ATCTTTACTT CAACAGAGAC
      721  GACGTGATCC TCTGTGTCTT GCCTATGTTC CATATATACG CTCTCAACTC CATCATGCTC
      781  TGTAGTCTCA GAGTTGGTGC CACGATCTTG ATAATGCCTA AGTTCGAAAT CACTCTCTTG
      841  TTAGAGCAGA TACAAAGGTG TAAAGTCACG GTGGCTATGG TCGTGCCACC GATCGTTTTA
      901  GCTATCGCGA AGTCGCCGGA GACGGAGAAG TATGATCTGA GCTCGGTTAG GATGGTTAAG
      961  TCTGGAGCAG CTCCTCTTGG TAAGGAGCTT GAAGATGCTA TTAGTGCTAA GTTTCCTAAC
     1021  GCCAAGCTAG GTCAGGGCTA TGGGATGACA GAAGCAGGTC CGGTGCTAGC AATGTCGTTA
     1081  GGGTTTGCTA AAGAGCCGTT TCCAGTGAAG TCAGGAGCAT GTGGTACGGT GGTGAGGAAC
     1141  GCCGAGATGA AGATACTTGA TCCAGACACA GGAGATTCTT TGCCTAGGAA CAAACCCGGC
     1201  GAAATATGCA TCCGTGGCAA CCAAATCATG AAAGGCTATC TCAATGACCC CTTGGCCACG
     1261  GCATCGACGA TCGATAAAGA TGGTTGGCTT CACACTGGAG ACGTCGGATT TATCGATGAT
     1321  GACGACGAGC TTTTCATTGT GGATAGATTG AAAGAACTCA TCAAGTACAA AGGATTTCAA
     1381  GTGGCTCCAG CTGAGCTAGA GTCTCTCCTC ATAGGTCATC CAGAAATCAA TGATGTTGCT
     1441  GTCGTCGCCA TGAAGGAAGA AGATGCTGGT GAGGTTCCTG TTGCGTTTGT GGTGAGATCG
     1501  AAAGATTCAA ATATATCCGA AGATGAAATC AAGCAATTCG TGTCAAAACA GGTTGTGTTT
     1561  TATAAGAGAA TCAACAAAGT GTTCTTCACT GACTCTATTC CTAAAGCTCC ATCAGGGAAG
     1621  ATATTGAGGA AGGATCTAAG AGCAAGACTA GCAAATGGAT TAATGAACTA G
//
LOCUS       HaCHS             1173 bp

        1  ATGGTTACTG TTGAAGAAGT TAGAAAAGCT CAAAGGGCAG AAGGTCCAGC CACAGTGATG
       61  GCTATTGGAA CCGCAGTTCC TCCAAATTGT GTAGATCAGG CCACTTATCC TGACTACTAC
```

```
     121  TTTAGAATAA CAAACTCTGA GCATAAGGCT GAATTGAAAG AAAAGTTCCA AAGGATGTGC
     181  GACAAATCAC AGATCAAGAA AAGATACATG TACCTTAATG AGGAAGTCCT AAAGGAAAAC
     241  CCAAATATGT GTGCATACAT GGCCCCTTCC CTTGACGCTA GACAAGATAT TGTGGTTGTA
     301  GAGGTCCCAA AATTGGGCAA GGAAGCAGCT GTTAAAGCCA TAAAGGAATG GGGTCAACCT
     361  AAGAGCAAAA TCACCCACCT TGTGTTTTGC ACTACAAGCG GAGTTGACAT GCCAGGCGCA
     421  GATTATCAGC TAACCAAACT TTTGGGTTTA AGGCCTTCTG TAAAAAGATT GATGATGTAC
     481  CAACAAGGTT GTTTCGCTGG AGGCACTGTC TTAAGACTAG CCAAGGATCT TGCAGAGAAC
     541  AACAAAGGTG CTAGGGTGTT GGTTGTATGC TCAGAAATTA CAGCCGTCAC CTTTAGAGGA
     601  CCAACTGACA CTCACTTAGA TTCCCTAGTT GGTCAGGCAT TGTTTGGCGA CGGTGCTGCC
     661  GCAATAATCA TTGGAAGTGA TCCTATTCCA GAGGTGGAAA AGCCTCTTTT TGAACTTGTT
     721  AGCGCTGCCC AAACTATATT GCCAGATTCT GAGGGTGCAA TCGACGGCCA CTTAAGGGAA
     781  GTAGGTCTAA CCTTCCATCT TTTGAAAGAT GTCCCTGGTT TAATTTCAAA GAACGTGGAA
     841  AAATCCCTAA CAGAGGCTTT TAAACCATTG GGTATAAGTG ACTGGAATAG CTTATTCTGG
     901  ATCGCTCACC CAGGCGGCCC TGCCATACTT GACCAGGTTG AAGCAAAATT GAGCTTAAAG
     961  CCAGAAAAAC TAAGAGCTAC TAGACATGTA TTGTCAGAGT ATGGTAACAT GTCCAGTGCC
    1021  TGTGTCCTTT TCATTTTGGA TGAAATGAGG AGAAAAAGCA AGGAGGACGG CCTAAAAACC
    1081  ACAGGTGAGG GAATCGAATG GGGTGTTCTA TTCGGCTTTG GTCCAGGCCT TACTGTGGAG
    1141  ACAGTTGTAC TTCATTCAGT CGCAATTAAT TAG
//
LOCUS       PhCHI       726 bp

       1 ATGTCTCCAC CAGTTTCTGT TACAAAAATG CAAGTCGAAA ATTATGCTTT TGCACCAACA
      61 GTGAACCCTG CCGGTTCCAC CAATACTTTG TTCTTAGCTG GAGCAGGCCA TAGAGGTCTA
     121 GAGATTGAAG GAAAGTTTGT GAAATTCACA GCCATAGGCG TATACCTTGA GGAAAGTGCT
     181 ATCCCATTTT TGGCAGAAAA GTGGAAAGGT AAGACCCCTC AGGAGTTAAC TGATAGCGTC
     241 GAGTTCTTTA GGGACGTGGT TACAGGTCCA TTCGAAAAGT TTACCAGAGT AACTATGATT
     301 CTACCTCTTA CAGGAAAGCA ATATTCTGAG AAAGTCGCCG AAAACTGTGT TGCTCACTGG
     361 AAGGGCATAG GTACCTACAC TGATGACGAA GGAAGGGCAA TCGAGAAATT CTTGGATGTG
     421 TTTAGATCAG AAACATTCCC ACCTGGTGCT TCCATTATGT TTACTCAGAG TCCATTAGGC
     481 TTGTTAACCA TCAGCTTTGC CAAGGACGAT TCAGTTACCG GTACTGCAAA TGCTGTAATC
     541 GAGAACAAAC AACTATCAGA AGCCGTCCTT GAATCCATTA TTGGAAAGCA TGGTGTGAGT
     601 CCTGCAGCCA AATGCTCTGT TGCCGAGAGA GTAGCAGAAT TGTTAAAAAA GAGCTATGCT
     661 GAAGAGGCCT CAGTGTTCGG CAAACCAGAA ACCGAAAAGT CCACAATACC TGTTATCGGT
     721 GTGTAG
//
```

Table 5. List of oligonucleotides		
CCM pathway genes and promoters		
Primer name	**Primer sequence, 5´ to 3´**	
ID1564_PTEF1->_U2_fw	CGTGCGAU GCACACACCATAGCTTC	Forward primer for USER cloning of the TEF1 promoter
ID1565_PTEF1->_U2_rv	ATGACAGAU TTGTAATTAAAACTTAG	Reverse primer for USER cloning of the TEF1 promoter
ID3108_PTEF1 _for_fusion_fw	AGCTACTGAU GCACACACCATAGCTTC	Forward primer for USER cloning of the TEF1 promoter fused with the TDH3 promoter for bidirectional expression
ID3107_PTDH3 _for_fusion_fw	ATCAGTAGCU ATAAAAAACACGCTTTTTCAG	Forward primer for USER cloning of the TDH3 promoter fused with the TEF1 promoter for bidirectional expression
ID1853_PTDH3 <-_U1_rv	ACCTGCACU TTTGTTTGTTTATGTGTGTTTATT C	Reverse primer for USER cloning of the TDH3 promoter
ID3097_CaCat A_U2_fw	ATCTGTCAU AAAACAATGTCCCAAG	Forward primer for USER cloning of CaCatA
ID3098_CaCat A_U2_rv	CACGCGAU TTACAACTTGATTTCAGC	Reverse primer for USER cloning of CaCatA
ID3103_KpAro Y.B_U1_fw	ATCTGTCAU AAAACAATGATCTGTCC	Forward primer for USER cloning of KpAroY.B
ID3104_KpAro Y.B _U1_rv	CGTGCGAU TCATTCAATTTCTTGAGC	Reverse primer for USER cloning ofKpAroY.B
ID3105_KpAro Y.Ciso _U2_fw	ATCTGTCAU AAAACAATGACCGCCCCAATC	Forward primer for USER cloning of KpAroY.Ciso
ID3016_KpAro Y.Ciso _U2_rv	CACGCGAU TTACTTAGCGGAACCTTGATTC	Reverse primer for USER cloning of KpAroY.Ciso
ID3095_KpAro Y.D _U2_fw	ATCTGTCAU AAAACAATGATCTGTCC	Forward primer for USER cloning of KpAroY.D
ID3096_KpAro Y.D _U2_rv	CACGCGAU TTATCTCTTATCTTCTGG	Reverse primer for USER cloning of KpAroY.D
ID3101_PaAro Z_U1_fw	AGTGCAGGU AAAACAATGCCATCCAAG	Forward primer for USER cloning of PaAroZ
ID3102_PaAro Z _U1_rv	CGTGCGAU TCACAAAGCAGCTGACAAAG	Reverse primer for USER cloning of PaAroZ
ID1391_ScTkl1 _U1_fw	AGTGCAGGU AAAACAATGACTCAATTCACTGA CATTG	Forward primer for USER cloning of Tkl1
ID1392_ScTkl1 _U1_rv	CGTGCGAU TCAGAAAGCTTTTTTCAAAGGAG	Reverse primer for USER cloning of Tkl1
LTTR sensor and reporter promoters		
Primer name	**Primer sequence, 5´ to 3´**	
MeLS069-F	GATGAATGCGGCCGCTTTA	Forward primer for random mutagenesis of BenM-EBD
MeLS093-R	CAATACGCCATCAAGTTGCTAAG C	Reverse primer for random mutagenesis of BenM-EBD
MeLS071-F	CTCCTTCCTTTTCGGTTAGAGCG GATGAATGCGGCCGCTTTA	Tailed primer for BenM-EBD library assembly by gap repair
MeLS094-R	TCATTTCTTTTACCAATACGCCAT CAAGTTGCTAAGC	Tailed primer for BenM-EBD library assembly by gap repair

MeLS001_F	ATCTGTCAUAAAACAATGGAATT GAGACAC	Forward primer for USER cloning of BenM
MeLS003_R	CACGCGAUTTACCAATTTGGTGG TTCAG	Reverse primer for USER cloning of BenM
MeLS005_F	CGTGCGAUATACTCCATAGGTAT TTT	Forward primer for USER cloning of the BenM binding site
MeLS008_R	CACGCGAUTTATTTGTACAATTC ATCCA	Reverse primer for USER cloning of the yeGFP ORF
MeLS009_F	ATCTGTCAUAAAACAATGTCTAA AGGTG	Forward primer for USER cloning of the yeGFP ORF
MeLS0046_F	CGTGCGAUTTCTTAGGCACAACA ATATTTATAAAAGAAG	Forward primer for USER cloning of the REV1 promoter
MeLS0047_R	ATGACAGAUCGCTGGATATGCC TAGAAATGC	Reverse primer for USER cloning of the REV1 promoter
MeLS0048_F	CGTGCGAUGGAAAACCAAGAAA TGAATTATATTTCC	Forward primer for USER cloning of the 491bp CYC1 promoter
MeLS0049_R	ATGACAGAUTATTAATTTAGTGT GTGTATTTGTGTTTGTG	Reverse primer for USER cloning of the CYC1 promoter
MeLS0052_F	CGTGCGAUCCAGGCAACTTTAG TGCTGACAC	Forward primer for USER cloning of the 209bp CYC1 promoter
MeLS0056_F	CGTGCGAUATAAAAAACACGCTT TTTCAGTTCG	Forward primer for USER cloning of the TDH3 promoter
MeLS0057_R	ATGACAGAUTTTGTTTGTTTATG TGTGTTTATTCGA	Reverse primer for USER cloning of the TDH3 promoter
MeLS0062_F	CGTGCGAUGAAAGACCACACCC ACGCG	Forward primer for USER cloning of the RNR2 promoter
MeLS0063_R	ATGACAGAUGGTAATTGGACAA ATAAATACGTGTATTAAG	Reverse primer for USER cloning of the RNR2 promoter
MeLS0064_F	CGTGCGAUTCAAGCCCACGCGT AGGC	Forward primer for USER cloning of the 272bp CYC1 promoter
MeLS0074_F	CGTGCGAUTCGAGCAGATCCGC CAGG	Forward primer for USER cloning of the 249bp CYC1 promoter
MelS101-R	CACGCGAUTCAGCAGATGCCTG GCAGC	Reverse primer for USER cloning of FdeR
MelS108-F	ATCTGTCAUAAAACAATGCGTTT CAACAAGCTCGAC	Forward primer for USER cloning of FdeR
TISNO-53F	ATCTGTCAUAAAACAATGATTGA TGCACGT	Forward primer for USER cloning of PcaQ
TISNO-54R	CACGCGAUTTATGCTGTTCTTTT GGCTTC	Reverse primer for USER cloning of PcaQ
TISNO-57F	ATCTGTCAUAAAACAATGAAACG TCCTGA	Forward primer for USER cloning of ArgP
TISNO-58R	CACGCGAUTCAATCTTGTCTTAA CACCTTATG	Reverse primer for USER cloning of ArgP
TISNO-59F	ATCTGTCAUAAAACAATGAAGGA CGACAT	Forward primer for USER cloning of MdcR
TISNO-60R	CACGCGAUTTACCTGTTAATAGA ACGAGCATA	Reverse primer for USER cloning of MdcR
Genotyping primers		
Primer name	**Primer sequence, 5´ to 3´**	
ID893 XII-2-up-out-sq	CGAAGAAGGCCTGCAATTC	Genotyping of genomic integration locus
ID894 XII-2-down-out-sq	GGCCCTGATAAGGTTGTTG	Genotyping of genomic integration locus
ID897 XII-4-	GAACTGACGTCGAAGGCTCT	Genotyping of genomic integration locus

out-seq_fw		
ID898 XII-4-down-out-sq	CGTGAAATCTCTTTGCGGTAG	Genotyping of genomic integration locus
ID903_X-3-out-seq_fw	TGACGAATCGTTAGGCACAG	Genotyping of genomic integration locus
ID904 X-3-down-out-sq	CCGTGCAATACCAAAATCG	Genotyping of genomic integration locus
ID905_X-4-out-seq_fw	CTCACAAAGGGACGAATCCT	Genotyping of genomic integration locus
ID906 X-4-down-out-sq	GACGGTACGTTGACCAGAG	Genotyping of genomic integration locus
ID907_XI-1-out-seq_fw	CTTAATGGGTAGTGCTTGACACG	Genotyping of genomic integration locus
ID908 XI-1-down-out-sq	GAAGACCCATGGTTCCAAGGA	Genotyping of genomic integration locus
ID911_XI-3-out-seq_fw	GTGCTTGATTTGCGTCATTC	Genotyping of genomic integration locus
ID2220 Sc_ColoPCR_fw	CCTGCAGGACTAGTGCTGAG	Genotyping of genomic integration locus
ID2221 Sc_ColoPCR_rv	GTTGACACTTCTAAATAAGCGAATTTC	Genotyping of genomic integration locus
MeLS0082_F	AAAAATAAATAGGGACCTAGACTTCAGG	Sequencing primer located in the CYC1 terminator
MeLS0053_R	CTGCAGGAATTCGATATCAAGC	Reverse sequencing primer located in the KI.URA terminator
MeLS0054_F	TCAATTGAGATGAGCTTAATCATGTC	Forward sequencing primer located in the KI.URA promoter
MeLS0055_R	ATTATTACAGTCACTCAGACAGAGCAC	Sequencing primer located in the XII-1 down homology region
MeLS0058_F	GTGAAGTGATCATGCACATCGC	Sequencing primer located in the REV1 promoter
MeLS0083_R	GAGGTTCCAGACCAGTTAAGACTACTC	Sequencing primer located in BenM DNA-binding domain
TISNO-15F	GTAAGCCAGATTAAAATTCACG	Sequencing primer located in REV1 promoter
TISNO-62R	TAGCATCACCTTCACCTTCACC	Genotyping of genomic integration locus (anneals to yeGFP)
TISNO-63F	CCTGAAATTATTCCCCTACTTGAC	Sequencing primer located in TDH3 promoter
TISNO-65R	TATCGGATAACAACACCGCTG	Genotyping of genomic integration locus (anneals to REV1p)
TISNO-66R	CTGTTCACCCAGACACCTAC	Genotyping of genomic integration locus (anneals to TDH3p)
TISNO-67R	GCGGAGTCCGAGAAAATCTG	Genotyping of genomic integration locus (anneals to TEF1p)
Res417 R	TCTCAGGTATAGCATGAGGTCGCTCAT	Genotyping of genomic integration locus, internal assembler primer
Res418 F	CCTGCAGGACTAGTGCTGAGGCATTAAT	Genotyping of genomic integration locus, internal assembler primer
RES395 XI-2 UP F	GTTTGTAGTTGGCGGTGGAG	Genotyping of genomic integration locus
RES396 XI-2 DW	GAGACAAGATGGGGCAAGAC	Genotyping of genomic integration locus
RES511 XVI-20 UP F	GGCTTGTGGTCACCTGTCAT	Genotyping of genomic integration locus
RES512 XVI-20 DW R	GAATTATGGTAATTTTGATTATC	Genotyping of genomic integration locus

RES658 X-2-UP F	TGCGACAGAAGAAAGGGAAG	Genotyping of genomic integration locus
RES659 X-2 DW R	GAGAACGAGAGGACCCAACAT	Genotyping of genomic integration locus
Naringenin pathway genes and promoters		
Primer name	**Primer sequence, 5´ to 3´**	
RES194 C4H F	AGCGATACGUAAAATGGATTTGTTATTGCTGGAAAAG	Forward primer for USER cloning of the C4H
RES195 C4H R	CACGCGAUTTACCATACATCTCTCAGATATCTAC	Reverse primer for USER cloning of the C4H
RES196 AtPAL2 F	ATCAACGGGUAAAATGGACCAAATTGAAGCAATGC	Forward primer for USER cloning of the AtPAL2
RES197 AtPAL2 R	CGTGCGAUTTAGCAGATTGGAATAGGTGCAC	Reverse primer for USER cloning of the AtPAL2
RES198 At4CL2 F	AGCGATACGUAAAATGACGACACAAGATGTGATAGTC	Forward primer for USER cloning of the 4Cl2
RES199 At4CL2 R	CACGCGAUCTAGTTCATTAATCCATTTGCTAG	Reverse primer for USER cloning of the 4Cl2
RES569 CHI F	AGCGATACGUAAAATGTCTCCACCAGTTTCTGTTAC	Forward primer for USER cloning of the CHI
RES570 CHI R	CACGCGAUCTACACACCGATAACAGGTATTG	Reverse primer for USER cloning of the CHI
RES571 CHS F	CGTGCGAUTAATTAATTGCGACTGAATGAAG	Forward primer for USER cloning of the CHS
RES572 CHS R	ATCAACGGGUAAAATGGTTACTGTTGAAGAAGTTAG	Reverse primer for USER cloning of the CHS
RES573 ATR2 L5 F	agctgcagcUaaagaagctgcagcaaaagctTCCAGTAGCTCTTCCTCCTC	Forward primer for USER cloning of the ATR2 (includes L5 linker)
RES 574 L5 C4H R	AgctgcagcUtcttttgctgcagcttcagcgctACAATTTCTGGGTTTCATG	Reverse primer for USER cloning of the C4H (includes L5 linker)
RES407 pTDH3 F	ACCCGTTGAUTTTTGTTTGTTTATGTGTGTTTATTCG	Forward primer for USER cloning of the TDH3 promoter
RES408 pTDH3 R	CACGCGAUGATCTCAGTTCGAGTTTATCATTATCA	Reverse primer for USER cloning of the TDH3 promoter
RES454 pPGK1 R	ACCCGTTGAUGCCGCTTGTTTTATATTTGTTGTAAAAAG	Reverse primer for USER cloning of the PGK1 promoter
RES455 pPGK1 F	CACGCGAUGGCCTGGAAGTACCTTCAAAGAATG	Forward primer for USER cloning of the PGK1 promoter
RES456 pTEF1 F	CGTGCGAUGCCGCACACACCATAGCTTCAAAATG	Forward primer for USER cloning of the TEF1 promoter
RES457 pTEF1 R	ACGTATCGCUGTGAGTCGTATTACGGATCC	Reverse primer for USER cloning of the TEF1 promoter
RES568 pPDC1 F	CGTGCGAUGCCGATCTATGCGACTGGGTGAG	Forward primer for USER cloning of the PDC1 promoter
RES640 pPDC1 R	ACGTATCGCUTTTTGATAGATTTGACTGTGTTATTTTGCG	Reverse primer for USER cloning of the PDC1 promoter
RES460 pTDH3/pTEF2 R	ACCCGTTGAUTTTTGTTTGTTTATGTGTG	Reverse primer for USER cloning of the bidirectional promoter
RES461 pTDH3 /pTEF2 F	ACGTATCGCUTGTTTAGTTAATTATAGTTC	Forward primer for USER cloning of the bidirectional promoter

Table 6

AlsR	Activator (local)	*Bacillus subtilis*	Renna *et al.* (1993) Bartowsky
AmpR	Activator (local)	*Rhodobacter*	& Normark (1993)
	Activator (local)	*capsulatus*	
ArgP	Activator (global)	*Enterobacter cloacae*	Nandineni & Gowrishankar
BenM	Activator (local)	*Citrobacter freundii*	(2004) Collier *et al.* (1998)
BlaA	Activator (global)	*Escherichia*	Raskin *et al.* (2003)
CatM	Activator (global)	*Acinetobacter* spp.	Chugani *et al.* (1998) van
CbbR	Activator (global)	*Streptomyces* spp.	Keulen *et al.* (2003)
CfxR		*Acinetobacter calcolaceticus* *Pseudomonas putida* *Xanthobacter flavus*	Windho¨vel 991)
ChiR	Activator (local)	*Serratia marcescens*	Suzuki *et al.* (2001)
CidR	Activator (local)	*Staphylococcus* spp.	Yang *et al.* (2005)
		Bacillus anthracis	Ahn *et al.* (2006)
ClcR	Activator (local)	*Pseudomonas putida*	Coco *et al.* (1993)
CrgA	Activator/Repressor (global)	*Neisseria meningitidis*	Deghmane *et al.* (2000)
CynR	Activator (local)	*Escherichia coli*	Sung & Fuchs (1992)
CysB	Activator (global)	*Salmonella enterica* Typhimurium *Escherichia coli*	van der Ploeg *et al.* (1997)
CysL	Activator (global)	*Bacillus subtilis*	Guillouard *et al.* (2002)
GltC	Activator (local)	*Bacillus subtilis*	Picossi *et al.* (2007)
HupR	Activator (global)	*Vibrio vulnificus*	Litwin & Quackenbush (2001)
HvrB	Activator (global)	*Rhodobacter*	Bugay *et al.* (1994)
IlvR	Activator (local)	*Caulobacter*	Malakooti & Ely (1994)
IlvY	Activator (local)	*Escherichia coli*	Wek & Hatfield (1988)
IrgB	Activator (local)	*Vibrio cholerae*	Goldberg *et al.* (1991)
LeuO	Activator/Repressor (global)	*Salmonella enterica* Typhimurium	Herna´ndez-Lucas *et al.* (2008)
LrhA	Activator (global)	*Escherichia coli*	Lehnen *et al.* (2002)
LysR	Activator (local)	*Escherichia coli*	Stragier *et al.* (1983)
MdcR	Activator (local)	*Klebsiella pneumoniae*	Peng *et al.* (1999)
MetR	Activator (global)	*Streptococcus* spp.	Kovaleva & Gelfand (2007)
MleR	Activator (local)	*Lactococcus lactis*	Renault *et al.* (1989)
MtaR	Activator (global)	*Group B streptococci*	Shelver et al. (2003)
MvfR	Activator (global)	*Pseudomonas aeruginosa*	Cao *et al.* (2001)
NagR	Activator (local)	*Ralstonia eutropha*	Jones *et al.* (2003)
NahR	Activator (local)	NAH7 plasmid of	Park *et al.* (2002)
NhaR	Activator (local)	Escherichia coli	Dover & Padan (2001)
NocR	Activator (local)	Ti plasmids of Agrobacterium	von Lintig et al. (1994)

Table 7

Super family	Specific ligand	Transcriptional activator	Operator sequence	reference
AraC/XylS	p-hydroxybenzoate	pobR	GCCGGCGCATGCGCCGCCGGCCAGCCATAA	http://www.pseudomonas.com/feature/intergenic?start=280759&stop=280935&repliconid=136&src=map
LuxR	N-(3-oxodo-decanoyl) homoserine lactone (3-oxo-C12-HSL) and N-butyryl homoserine lactone	LasR	CTATGTCTTTTGTTAG	http://www.pseudomonas.com/feature/intergenic?start=1077905&stop=1078461&repliconid=136&src=map
LTTR	multiple quorum sensing	mvfR	TTCGGACTCCGAA	http://www.pseudomonas.com/feature/intergenic?start=1077905&stop=1078461&repliconid=136&src=map + Xiao et al 2006
LTTR	flavonoids, naringenin, hesperetin	nodD	AGATTAGTAAAATTGATTGTTGGGATAGCTATCATCCACGATATGGATG	yang et al 2012; http://www.ncbi.nlm.nih.gov/gene/4403938
	benzoate	benR	CCGAAAAAGTACCGAACATCCGTAAATCTGGATAACGTTCTGCACAATCCGGATAGCCCCCCGCCAGCCGTCTCCCTAAC	putativ, COWLESet al 2000, http://www.ncbi.nlm.nih.gov/gene/1046807
Lrp/AsnC	binding inhibition with lysine	lysM	TAAAATCGTACCACTTATTACTAAAAACTTTTTCTACACAAAACTAAGTTAGTATCTAAC	Brinkman et al 2002+http://www.ncbi.nlm.nih.gov/gene/1453332
LTTR	sulfur sources? (no cysteine) N-acetyl-serine	cysB	TGTTGAAATTAAAGGCCTTTAGAAACTTGAATTCTATGGACCGAACTAAAA	Hryniewicz et al, 1991; Delic-Atree 1997 + http://www.pseudomonas.com/feature/show/?id=109899&view=sequence
LTTR	sulfur sources? (no cysteine) N-acetyl-serine	cysBH		
LTTR	muconic acid	benM	See Fig. 8	

LTTR	naringenin	FdeR	See Fig. 8	Siedler S et al., 2014
LTTR	salicylate	NagR		Jones et al., 2003
LTTR	salicylate	NahR		Cebolla et al, 1997; van Sint Fiet et al, 2006; Calcagno review
LTTR	protcatechuic acid	PcaQ	See Fig. 8	MacLean et al, 2008
LTTR	acetate	AlsR		Fradrich et al, 2013, de Oliveira et al, 2013
LTTR	L-arginin	ArgP	See Fig. 8	Zhou et al, 2010; Laisram et al., 1997
LTTR	malonate	MdcR	See Fig. 8	
LTTR		AphB	AACAACCTA AGTTTGCA	Kovacikova et 2010, Bina et al 2015
	ROS, superoxide, paraquate	soxS	TTTGCATAG CGTGAATAT GTCAAAATT GAT	Gil et al 2009+http://www.ncbi.nlm.nih.gov/gene/1253251
LTTR	myricetin and kaempferol	kaeR	**CGATTTGC CATTAATC CCATTAGG ACTTTCGT ATCGGA**GA AGCCTTCAA CGTTATTAA ACATCATTG CTGGACCTT CTTGCGTCG GCCGTTTTA CCGTCCCTC CAGCACCAA TATAGCGGT AAACACCAG CCAATTCAG CATTTGGAT TCACAGCTA CGTTCGTCT CATGGTACT GGTTGGCA TGGGTTTTT AGCTCGGC CAATACTTT TCGTAAATC ATAAGGATC ATTTACCAT CAGATTACC TCCTATAAG TTGCTTACA ATCACCACT TTAAGGCAT AAAATCGTT GCAAACAAC TCAACTTTC GACTAA**TG TTATGCCT AAATGGAA TAATAAGA AGAAGGTT CTTCAAT**	Pande et al 2001; http://www.ncbi.nlm.nih.gov/gene/?term=kaeR

			(5')*	
	L-arabinose	araC	TCAGGCAG GATCCGCTA ATCTTATGG ACAAAAATG CTAATGCTT TGCAAAGT GTGACGCT GTGCAAATA TTCAATGTG GACATTCCA GCCATAGTT ATAGACACT TCTGTTACT TAATTTTAT CGCCTGAA CTGTACGCT TTTGTTACA AAGCGCTTT TCACAAGC GGGGTTGA TACGTGCTT TCATCAAGC GCAAAGTCT TGCGGAGA CGGAAGCT CTGTCGTCC TGGTCGATA TGGACAATT TGTTTC	http://www.ecogene.org/gene/EG10054, http://www.ncbi.nlm.nih.gov/gene/1251622

* Bold sequences are DNA sites where KaeR bind. The site is palindromic meaning that it will bind to two somewhat complementary motifs (the two bold sequences). The sequence in between indicates space which is not necessarily needed to be matching this code in length nor sequence content.

RESULTS

Onboarding a prokaryote transcription activator to yeast

To investigate the potential to build orthogonal biosensors using prokaryotic transcriptional activators in a eukaryotic chassis, we initially selected BenM from *Acinetobacter* sp ADP1 for several reasons. First, it belongs to the LTTR family, which is one of the most abundant families of transcriptional regulators found in a diverse range of prokaryotes. Second, in *Acinetobacter* sp ADP1, BenM serves as a native CCM-inducible transcriptional activator (Results, Fig. 1). CCM is an intermediate from aromatic compound catabolism and an important precursor for bioplastics. Moreover CCM biosynthesis was recently refactored in

yeast, yet without any high-throughput screening option available. Third, BenM has a well-characterized DNA binding site (herein termed *BenO*) and mode-of-action (Fig. 1a and Fig. 1). Finally, this protein does not require any binding to regulatory subunits apart from its cognate inducers, which should ensure its orthogonality in non-native chassis.

Engineering transcriptional repressors from prokaryotes into eukaryote chassis has emphasized the importance of operator positioning within synthetic eukaryote promoters in relation to transcriptional output. Hence, we first sought to identify optimal positioning of *BenO* when introduced into a eukaryote promoter. As a first expression cassette the full-length (491bp) CYC1 promoter (CYC1p) was used to control the expression of green fluorescence protein (GFP) (Olesen, J., Hahn, S. & Guarente, L. Yeast HAP2 and HAP3 activators both bind to the CYC1 upstream activation site, UAS2, in an interdependent manner. Cell 51, 953–61 (1987)). CYC1p was recently reported as a suitable promoter for introduction of other non-native TF binding sites in yeast, and throughout this study all engineered reporter gene promoters will be based on chromosomally integrated full-length or truncated versions of this promoter. Initially, *BenO* was introduced into the 491bp CYC1 promoter immediately upstream of one of the two TATA boxes - TATA-1β (designated 491bp_CYC1p_ *BenO_T1*) or TATA-2α (designated 491bp_CYC1p_*BenO_T2*), or upstream of both (designated 491bp_CYC1p_*BenO_T1/T2*)(Fig. 2a). Outputs from these engineered promoters were compared by flow cytometry to expression from the native CYC1p (491bp_CYC1p) using GFP as the reporter (Fig. 1a and Fig. 2a). In general, introducing *BenO* negatively impacted the CYC1p activity (Fig. 1b, white columns). However, when co-expressing BenM from the TEF1 promoter we observed 20-fold and 5-fold induction of expression from 491bp_CYC1p_*BenO_T1* and 491bp_CYC1p_*BenO_T1/T2* compared to the promoter activities without co-expression of BenM. For 491bp_CYC1p_*BenO_T2* we observed a modest 30% reduction in expression. Most importantly, BenM did not increase expression of native CYC1p without *BenO* (Fig. 1b). Taken together these data show that BenM can function as a transcriptional activator in yeast.

Protonated CCM is directly taken up by yeast at pH 4.5 without any growth defects (Fig. 3a-b). This enables CCM inducibility of the genetic devices to be tested by simple supplement of 200 mg/L CCM to the medium at pH 4.5. Following 24 h of cultivation GFP output was measured using flow cytometry. Here, we observed modest increases (1.3-2.2-fold, Fig. 1b) in reporter output from all CYC1 promoters that harbored *BenO*, whereas no change was observed from the native CYC1p (Fig. 1b). Also, all engineered promoters showed significant transcriptional activities in the control medium (no CCM) compared to background auto-fluorescence (Fig. 1b).

In order to lower the basal activity of the engineered promoters, we removed upstream activating sequences (UAS1 and UAS2) and introduced *BenO* into truncated versions of the CYC1p (designated 272bp_CYC1p, 249bp_CYC1p and 209bp_CYC1p, Fig. 2a). Also, in order to improve the dynamic range of the genetic device we tuned the production of BenM by placing *benM* under the transcriptional control of three other native yeast promoters: TDH3p, RNR2p and REV1p. Together with TEF1p, this system allows for an expression range covering almost three orders of magnitude. By combining and chromosomally integrating all possible BenM expression cassettes with all CYC1p-derived reporter constructs, a total of 84 yeast strains were generated, including control strains (Fig. 1c, Table 1 and Fig. 4). Analyzing basal and CCM-induced GFP expression for all strains by flow cytometry we observed reporter outputs that spanned more than two orders of magnitude from the lowest to the highest GFP levels, with most of the high outputs resulting from reporters expressed from full-length CYC1p backbones co-expressed with BenM (Fig. 1c, Fig. 4 and Table 2). Low-expressing strains mostly comprise truncated CYC1p reporter variants without *BenO* or BenM. These data showed that the *BenO_T1* positioning allowed CCM-inducibility of all truncated variants of CYC1p, with the highest dynamic range observed for the minimal promoters 249bp_CYC1p_*BenO*_T1 and 209bp_CYC1p_*BenO*_T1 (3.2-4.7-fold)(Fig. 1d and Table 2). Among the genetic devices tested, strain MeLS0049 with 209bp_CYC1p_*BenO*_T1 controlled by BenM expressed from REV1p showed both low basal activity and high CCM-inducibility (3.8-fold), and was therefore regarded as most suitable for application as a CCM biosensor.

High-throughput prototyping of biosensors variants

The dynamic range of a biosensor output is an important parameter when evaluating applicability of a biosensor for screening and selection. For this reason, we applied a high-throughput engineering strategy for identifying BenM mutants with higher dynamic ranges when expressed from the weak REV1 promoter. Previous mutagenesis studies identified residues important for ligand-binding in LTTR effector binding domains (EBDs). For this purpose we performed PCR-based mutagenesis of the BenM EBD (residues 90-304)(Fig. 2a). Following mutagenesis we harnessed yeast's homologous recombination machinery for plasmid gap repair of variant EBDs with the BenM DNA-binding domain (DBD)(Fig. 2a). A population derived from approx. 40,000 transformants was analyzed by fluorescence activated cell sorting (FACS) using a two-step approach, in which we first removed the variants showing increased basal activity. Next, we compared fluorescence output from the population of transformants in control and CCM medium (Fig. 2b). From this, all cells showing higher fluorescence than the fluorescence observed in control medium were sorted (Fig. 2b). Sorted cells were subsequently cultivated as clones and validated by flow cytometry (Fig. 2c). Here we identified five BenM variants with higher dynamic ranges than wild-type BenM (Fig. 2c). Sequencing of the BenM variants identified a triple mutant with point mutations H110R,

F211V and Y286N in the BenM EBD (Fig. 2c). Plasmid-based expression of BenM$^{H110R, F211V, Y286N}$ showed doubled GFP output upon CCM induction (6-fold), compared to induction for the plasmid-based expression of wild-type BenM (Fig. 2c). Interestingly, the mutations in BenM$^{H110R, F211V, Y286N}$ were not positioned in the immediate vicinity of the CCM binding site (Fig. 2d). Similar to all other genetic devices engineered in this study, BenM$^{H110R, F211V, Y286N}$ was also integrated into the genome for stable expression.

LTTR-based biosensor specificity and orthogonality

To assess the potential application of the LTTR-based biosensor for CCM in yeast, we next investigated the specificity of BenM, as well as its potential impact on the host transcriptome. First, by testing a range of diacids supplied to the growth medium at pH 4.5 with identical molar concentrations to CCM (1.4 mM), we observed that among the diacids tested both BenM and BenM$^{H110R, F211V, Y286N}$ induce GFP expression specifically in response to CCM (Fig. 3a). Second, to test for transcriptional orthogonality of BenM$^{H110R, F211V, Y286N}$ in yeast, we used RNA-seq to quantify and compare the transcriptomes of cells with (MeLS0284) or without (MeLS0138) expression of BenM$^{H110R, F211V, Y286N}$. As the genetic device has low basal activity (Fig. 4 and Table 2) we analyzed yeast transcriptomes following 24 h cultivation in the presence of CCM. Here, we observed that the average GFP transcript abundance from strain MeLS0284 was approximately 27-fold higher compared to strain MeLS0138 (Fig. 3b, Fig. 5). Apart from genes encoding GFP and BenM, only one other gene encoding the Golgi-associated retrograde protein complex component *TCS3*, passed our stringent cut-off ($P < 0.05$, >2-fold) showing a modest decrease (2.3x) in expression level when BenM$^{H110R, F211V, Y286N}$ was expressed (Fig. 3b). We found no match to *BenO* in this gene's promoter (data not shown), suggesting that the minor transcriptome perturbations could be due to noise in RNA-seq measurements or indirect effects.

A design for onboarding LTTR-based biosensors in yeast

The genetic device developed in this study represents to the best of our knowledge the first example of transplanting a prokaryotic transcriptional activator into a eukaryotic chassis and successfully using it to activate gene expression without the need for modifying the protein beyond codon optimization. Acknowledging the vast numbers of transcriptional activators found among LTTR members, the optimal reporter promoter design (209bp_CYC1p_*BenO_T1*) could prove valid for other metabolic engineering and biotechnological applications. To test the generality of the biosensor design for onboarding

other small-molecule binding transcriptional activators as biosensors in yeast we selected four other candidates from the LTTR family; FdeR from *Herbaspirillum seropedicae,* PcaQ from *Sinorhizobium meliloti*, ArgP from *Escherichia coli*, and MdcR from *Klebsiella pneumonia*, with co-inducers naringenin, protocatechuic acid (PCA), L-arginine, and malonic acid, respectively. In this proof-of-principle study we selected the four candidates based on a minimal set of information, including knowledge about operator sequences, experimental evidence for ligand-inducible control of target operons, and their mode-of-action within native chassis (ie. activation, Fig. 1). Furthermore all of these metabolites can passively diffuse across the yeast plasma membrane, with the exception of malonic acid, which requires the expression of the dicarboxylic acid transporter MAE1 from *Schizosaccharomyces pombe*. For this purpose, the gene encoding MAE1 was integrated into cells expressing MdcR (Table 3). Based on this knowledge, and the aforementioned selection criteria, we directly replaced *BenO* located in the T1 position of the 209bp_CYC1p promoter with operator sequences for each of these LTTRs (Fig. 4a, Fig. 2a-b, Table 4). We first tested if expression of GFP could be activated upon low and high expression of individual LTTRs. From this, it was evident that all LTTRs were able to activate GFP expression from the 209bp_CYC1p_T1 promoter when the LTTR was expressed from the strong TDH3 promoter compared to yeast cells without expression of an LTTR (1.4x-8.1x), with BenM showing the strongest activation (8.1x)(Fig. 4a). Similarly, GFP expression could also be induced by ArgP when the weak REV1p promoter controlled expression of the LTTR (2.2x). This proves the broad applicability of the reporter promoter design, and that biosensor output is tunable depending on the expression level of the LTTR. Next, we tested if each LTTR could further induce GFP expression when its cognate inducer was supplied to the growth medium (Fig. 4b). For this purpose we prepared medium with either 1.4 mM CCM, 0.2 mM naringenin, 30 mM L-arginine, 1.4 mM PCA, or 10 mM malonic acid, as previously reported to be relevant concentrations in terms of bio-based production and microbial physiology. Here, in addition to BenM, ArgP was the only LTTR enabling a significant ligand-inducible increase in GFP expression when LTTR expression was controlled by REV1p (Figure 4b). However, when expressing LTTRs from the TDH3 promoter all LTTRs, except PcaQ, significantly increased GFP expression (1.4x-4.1x) when their cognate ligand was present in the cultivation medium (Fig. 4b). Taken together all tested LTTRs were able to activate expression of GFP when their operators were placed in the T1 position of the 209bp_CYC1p scaffold promoter (Table 4). Furthermore, just as for BenM, yeast expressing FdeR, ArgP and MdcR from the strong TDH3 promoter, were able to further induce GFP expression upon addition of their cognate inducers (Figure 4b).

Many of the characterized LTTRs regulate operons by binding prototypic LTTR box patterns 5′-T-N11-A-3′ and 5′-TTA-N7/8-GAA-3′. In addition to transcriptional orthogonality (Fig. 3b), we therefore further tested if individual LTTRs would cross-react with operators for another LTTR. For this purpose, we expressed LTTRs ArgP and MdcR together with the

209bp_CYC1p_T1 promoter with operators for MdcR (herein *MdcO*) or ArgP (herein *ArgO*) driving the expression of GFP. As controls we tested GFP expression from 209bp_CYC1p_T1 promoter with *MdcO* or *ArgO* without expression of LTTRs. Flow cytometry analysis showed specificity between LTTR transcriptional activators and their inferred operator (Fig. 4c). This is in agreement with another study on cross-reactivity between promoter and transcriptional regulators of the TetR family, and the fact that LTTR residues in both the conserved N-terminal DNA-binding domains and the divergent EBDs are important for DNA-binding.

***In vivo* application of LTTR-based biosensors in yeast**

Based on our engineering efforts and characterization of prokaryote LTTR-based biosensors imported into yeast, we next addressed whether such biosensors would support real-time monitoring of product accumulation *in vivo* and thereby potentially provide high-throughput screening assays of biocatalysts. To test this we selected CCM and naringenin, for which highest titers in shake-flask cultivated haploid yeast of approx. 1 mM (141 mg/L) and 0.2 mM (54 mg/L), respectively, have recently been reported. Also, these two products are of general interest to biotechnology with CCM being a platform chemical for the production of several valuable consumer bio-plastics, whereas naringenin belongs to a class of secondary metabolites called flavonoids with nutritional and agricultural value.

Before applying the biosensors for *in vivo* detection of these metabolites we first tested their operational range and induction kinetics. For BenM and $BenM^{H110R, F211V, Y286N}$, we observed a weakly sigmoidal input-output relationship between CCM concentration and GFP output following 24 h cultivation. For chromosomally integrated $BenM^{H110R, F211V, Y286N}$ and BenM, a maximum of 10- and 3.5-fold induction was reached in the presence of the highest soluble CCM concentrations (1.4 mM, 200 mg/L)(Fig. 5a). Interestingly, induction kinetics of BenM and $BenM^{H110R, F211V, Y286N}$ were similar. This is in line with BenM mutations likely not to be involved with direct binding of CCM (Fig. 2d), but rather alter BenM binding to DNA to support increased GFP expression.

Similarly, for FdeR we first tested naringenin sensitivity and operational range of the sensor. As for CCM, the operational range was only tested for concentrations of naringenin soluble in growth medium (ie. < 0.2 mM). Here, we observed that expression of FdeR controlled by the weak REV1 promoter did not support induction of GFP expression at any of the tested concentrations (Fig. 5b), yet when expression of FdeR was controlled by the strong TDH3 promoter a maximum 1.7-fold increase in GFP expression was observed following 24 h cultivation in the presence of 0.2 mM naringenin (Fig. 5b). Taken together, the operational

ranges of BenM and FdeR are within the ranges of reported CCM and naringenin production titers in yeast, and therefore could make them applicable for screening such biocatalysts.

Next, we transformed the CCM biosensor (209bp_CYC1p_*BenO*_T1::GFP and REV1p::BenM$^{H110R, F211V, Y286N}$) into a small library of six yeast strains engineered to produce CCM. CCM production with a final titer of 149 mg/L was recently reported in haploid yeast using a three-step heterologous pathway consisting of a *AroZ* homologue from *Podospora anserina* encoding dehydroshikimate dehydratase (*Pa*AroZ), the *AroY* gene from *Klebsiella pneumonia* encoding the multi-subunit protocatechuic acid decarboxylase (PCA-DC) and the *CatA* gene encoding catechol 1,2-dioxygenase from *Candida albicans* (*Ca*CatA) (Fig. 6a). From that study it was clear that PCA-DC was a rate-limiting step for flux through the upper part of the shikimate pathway towards CCM. It was also suggested that an increased supply of precursor towards erythrose-4-phosphate (E4P) could improve CCM production. For this reason we introduced single or multiple copies of different PCA-DC subunits from *K. pneumonia* and introduced no or one additional copy of transketolase (Tkl1) from *S. cerevisiae* (Fig. 6a). The six-membered CCM production strain library and a wild-type CCM null background strain were cultured individually. After 24 h of cultivation the medium was analyzed for CCM concentration using HPLC and the cells were analyzed by flow cytometry for GFP intensity measurements. Here, we observed a strong correlation (r=0.98) between GFP output and CCM production titers, spanning a range of 0.00016-1.39 mM (0.023-197.6 mg/L)(Fig. 6b). The highest titers were obtained in strain ST4245-2 with multiple TY integrations of AroY subunits B and C and Tkl1 (Fig. 6a-b). To further examine the performance of the CCM biosensor we monitored GFP output and CCM production titers following 72 h of cultivation. Here, GFP outputs were saturated at titers >1.41 mM (200 mg/L)(Fig. 6a-b). However, the strain that produced the most CCM after 72 h (3.03 mM, 430.8 mg/L) also produced the most CCM and had the highest fluorescence after 24 h, emphasizing the applicability of the CCM biosensor for screening high-producing strains during early stages of cultivation.

Finally, we transformed 209bp_CYC1p_*FdeO*_T1::GFP and TDH3p::FdeR into yeast strains with a 5-step heterologous naringenin pathway. For building a small library of naringenin producing strains, we chromosomally introduced either in single copy of the pathway (EVR1), or with one and two additional integrations of bottleneck enzymes (AtPAL-2 and HaCHS for EVR2; *At*PAL-2, *Ha*CHS, and *At*C4H:L5:*At*ATR2 for EVR3)(Fig. 6c, Table 1). Following 48 h of cultivation the medium was analyzed for naringenin concentration using UPLC and the cells were analyzed by flow cytometry for GFP intensity measurements. As observed for the CCM biosensor, the naringenin biosensor also had a strong correlation (r=0.96) between GFP output and naringenin titers, spanning a range of 0.094-0.184 mM (25.61-50.18 mg/L)(Fig.

6d), with the highest titer obtained in strain EVR3 containing two additional integrations of bottleneck enzymes on top of the full copy of the 5-step naringenin pathway. For the naringenin sensor we observed a poorer correlation between biosensor output and titers at 24 h (r=0.87) compared to our 48 h (r=0.96) measurements (Fig. 6c-d). However, just as for the CCM biosensor, the strain that produced the most naringenin at 48 h (0.184 mM, 50.18 mg/L) also produced the most naringenin (0.045 mM, 12.25 mg/L) and had the highest fluorescence at 24 h.

Taken together, the two applications of the LTTR-based biosensors suggest that simple expression of the LTTR and an engineered reporter promoter (209bp_CYC1p_T1::GFP) with an operator site in position T1 allows for direct transplantation of prokaryotic transcriptional activators as biosensors to screen for the best-performing biocatalysts. Interestingly, though some of the transcriptional activators used in this study derived from prokaryotes with growth optima at higher temperatures compared to yeast, BenM showed a higher dynamic range in output at 30°C compared to 37°C (Fig. 7), illustrating robustness of LTTR performance.

Application of biosensors in CHO cells

For testing the reporter promoter design with other promoter backbones AND in another eukaryotes, Chinese hamster ovary cells was transformed using the human cytomegalovirus promoter backbone (CMV) instead of the CYC1 promoter backbone used in yeast.

Just as in the case with the yeast design using CYC1 promoter as a backbone, the present inventors put the binding site (benO) for the prokaryotic transcriptional activator BenM 6 bp upstream of the TATA box the CMV promoter and scored reporter gene activity (GFP fluorescence) in the presence and absence of the transcriptional activator BenM. As can be seen in figure 13, the design worked when putting into another promoter backbone AND another host organism (CHO cells).

In addition to this, the present inventors also tested 17 different positions for positioning of the BenM binding site (benO). Only position 6 upstream of the TATA box gave a significant response (See figure 14).

Prokaryotic operator benO for the prokaryotic transcriptional activator BenM is placed 6 bp upstream of the TATA box the CMV promoter

```
1   GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA
51  TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA
```

101 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
151 TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA
201 TGGGTGGAGT ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA
251 TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT AAATGGCCCG
301 CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
351 TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA
401 GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
451 TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG
501 GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG
551 TAGGCGTGTA CGGT**GGATAC TCCATAGGTA TTTTATTATA CAAATAATGT**
601 **GTTTGAACTT ATTAAAACAT TCTTTTAAGG TATAAACAA**G AGGTC***TATA***T
651 AAGCAGAGCT **C*GCCACC***ATG

BenO (Bold)

TATA (Bold,Italic)

KOZAK (Bold,Italic, underlined)

START CODON (underlined)

LIST OF REFERENCES

1. Jakočiūnas, T., Jensen, M. K. & Keasling, J. D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **34,** 44–59 (2015).

2. Esvelt, K. M. & Wang, H. H. Genome-scale engineering for systems and synthetic biology. *Mol. Syst. Biol.* **9,** 641 (2013).

3. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–8 (2000).

4. Wang, B., Barahona, M. & Buck, M. Amplification of small molecule-inducible gene expression via tuning of intracellular receptor densities. *Nucleic Acids Res.* **43,** 1955–64 (2015).

5. Farzadfard, F. & Lu, T. K. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science (80-.).* **346,** 1256272–1256272 (2014).

6. Michener, J. K. & Smolke, C. D. High-throughput enzyme evolution in Saccharomyces cerevisiae using a synthetic RNA switch. *Metab. Eng.* **14,** 306–16 (2012).

7. Raman, S., Rogers, J. K., Taylor, N. D. & Church, G. M. Evolution-guided optimization of biosynthetic pathways. *Proc. Natl. Acad. Sci.* **111,** 201409523 (2014).

8. Choi, J. H. & Ostermeier, M. Rational design of a fusion protein to exhibit disulfide-mediated logic gate behavior. *ACS Synth. Biol.* **4,** 400–6 (2015).

9. Ausländer, S., Ausländer, D., Müller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocomputers. *Nature* **487,** 123–127 (2012).

10. Khalil, A. S. *et al.* A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150,** 647–58 (2012).

11. Folcher, M., Xie, M., Spinnler, A. & Fussenegger, M. Synthetic mammalian trigger-controlled bipartite transcription factors. *Nucleic Acids Res.* **41,** e134 (2013).

12. Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10,** 99–105 (2014).

13. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 5547–51 (1992).

14. Stanton, B. C. *et al.* Systematic transfer of prokaryotic sensors and circuits to mammalian cells. *ACS Synth. Biol.* **3,** 880–91 (2014).

15. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154,** 442–51 (2013).

16. Gossen, M. *et al.* Transcriptional activation by tetracyclines in mammalian cells. *Science* **268,** 1766–9 (1995).

17. Teo, W. S. & Chang, M. W. Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in Saccharomyces cerevisiae. *Biotechnol. J.* **10,** 315–22 (2015).

18. Lee, N., Francklyn, C. & Hamilton, E. P. Arabinose-induced binding of AraC protein to araI2 activates the araBAD operon promoter. *Proc. Natl. Acad. Sci. U. S. A.* **84,** 8814–8 (1987).

19. Shadel, G. S. & Baldwin, T. O. The Vibrio fischeri LuxR protein is capable of bidirectional stimulation of transcription and both positive and negative regulation of the luxR gene. *J. Bacteriol.* **173,** 568–74 (1991).

20. Lee, D. J., Minchin, S. D. & Busby, S. J. W. Activating Transcription in Bacteria. *Annu. Rev. Microbiol.* **66,** 125–52 (2012).

21. Siedler, S., Stahlhut, S. G., Malla, S., Maury, J. & Neves, A. R. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into Escherichia coli. *Metab. Eng.* **21,** 2–8 (2014).

22. Maddocks, S. E. & Oyston, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154,** 3609–23 (2008).

23. Collier, L. S., Gaines, G. L. & Neidle, E. L. Regulation of benzoate degradation in Acinetobacter sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J. Bacteriol.* **180,** 2493–501 (1998).

24. Suastegui, M. *et al.* Combining Metabolic Engineering and Electrocatalysis: Application to the Production of Polyamides from Sugar. *Angew. Chemie* **128,** 2414–2419 (2016).

25. Curran, K. A., Leavitt, J. M., Karim, A. S. & Alper, H. S. Metabolic engineering of muconic acid production in Saccharomyces cerevisiae. *Metab. Eng.* **15,** 55–66 (2013).

26. Bundy, B. M., Collier, L. S., Hoover, T. R. & Neidle, E. L. Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7693–8 (2002).

27. Wang, M., Li, S. & Zhao, H. Design and engineering of intracellular-metabolite-sensing/regulation gene circuits in Saccharomyces cerevisiae. *Biotechnol. Bioeng.* **113,** 206–15 (2016).

28. Olesen, J., Hahn, S. & Guarente, L. Yeast HAP2 and HAP3 activators both bind to the CYC1 upstream activation site, UAS2, in an interdependent manner. *Cell* **51,** 953–61 (1987).

29. McIsaac, R. S., Gibney, P. A., Chandran, S. S., Benjamin, K. R. & Botstein, D. Synthetic biology tools for programming gene expression without nutritional perturbations in Saccharomyces cerevisiae. *Nucleic Acids Res.* (2014). doi:10.1093/nar/gkt1402

30. Li, W. Z. & Sherman, F. Two types of TATA elements for the CYC1 gene of the yeast Saccharomyces cerevisiae. *Mol. Cell. Biol.* **11,** 666–76 (1991).

31. Pfeifer, K., Arcangioli, B. & Guarente, L. Yeast HAP1 activator competes with the factor RC2 for binding to the upstream activation site UAS1 of the CYC1 gene. *Cell* **49,** 9–18 (1987).

32. Lee, M. E., Aswani, A., Han, A. S., Tomlin, C. J. & Dueber, J. E. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.* **41,** 10668–78 (2013).

33. Peng, H. L., Shiou, S. R. & Chang, H. Y. Characterization of mdcR, a regulatory gene of the malonate catabolic system in Klebsiella pneumoniae. *J. Bacteriol.* **181,** 2302–6 (1999).

34. MacLean, A. M., MacPherson, G., Aneja, P. & Finan, T. M. Characterization of the beta-ketoadipate pathway in Sinorhizobium meliloti. *Appl. Environ. Microbiol.* **72,** 5403–13 (2006).

35. Laishram, R. S. & Gowrishankar, J. Environmental regulation operating at the promoter clearance step of bacterial transcription. *Genes Dev.* **21,** 1258–72 (2007).

36. Maclean, A. M., Haerty, W., Golding, G. B. & Finan, T. M. The LysR-type PcaQ protein regulates expression of a protocatechuate-inducible ABC-type transport system in Sinorhizobium meliloti. *Microbiology* **157,** 2522–33 (2011).

37. Chen, W. N. & Tan, K. Y. 'Malonate uptake and metabolism in Saccharomyces cerevisiae'. *Appl. Biochem. Biotechnol.* **171,** 44–62 (2013).

38. Opekarová, M. & Kubín, J. On the unidirectionality of arginine uptake in the yeast Saccharomyces cerevisiae. *FEMS Microbiol. Lett.* **152,** 261–7 (1997).

39. Rogers, J. K. & Church, G. M. Genetically encoded sensors enable real-time observation of metabolite production. *Proc. Natl. Acad. Sci. U. S. A.* (2016). doi:10.1073/pnas.1600375113

40. Rikhvanov, E. G., Varakina, N. N., Rusaleva, T. M., Rachenko, E. I. & Voĭnikov, V. K. [The effect of sodium malonate on yeast thermotolerance]. *Mikrobiologiia* **72,** 616–20

41. Koopman, F. *et al.* De novo production of the flavonoid naringenin in engineered Saccharomyces cerevisiae. *Microb. Cell Fact.* **11,** 155 (2012).

42. Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* **126,** 485–93 (2001).

43. Naesby, M. *et al.* Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in Saccharomyces cerevisiae. *Microb. Cell Fact.* **8,** 45 (2009).

44. Gupta, R. K., Patterson, S. S., Ripp, S., Simpson, M. L. & Sayler, G. S. Expression of the Photorhabdus luminescens lux genes (luxA, B, C, D, and E) in Saccharomyces cerevisiae. *FEMS Yeast Res.* **4,** 305–13 (2003).

45. Galloway, K. E., Franco, E. & Smolke, C. D. Dynamically reshaping signaling networks to program cell fate via genetic controllers. *Science* **341,** 1235005 (2013).

46. Kim, T., Folcher, M., Doaud-El Baba, M. & Fussenegger, M. A synthetic erectile optogenetic stimulator enabling blue-light-inducible penile erection. *Angew. Chem. Int. Ed. Engl.* **54,** 5933–8 (2015).

47. Zhang, H., Li, Z., Pereira, B. & Stephanopoulos, G. Engineering E. coli-E. coli cocultures for production of muconic acid from glycerol. *Microb. Cell Fact.* **14,** 134 (2015).

48. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7,** 562–78 (2012).

49. Jensen, N. B. *et al.* EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. *FEMS Yeast Res.* **14,** 238–48 (2014).

50. Gietz, R. D. & Schiestl, R. H. Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2,** 38–41 (2007).

51. Eckert-Boulet, N., Pedersen, M. L., Krogh, B. O. & Lisby, M. Optimization of ordered plasmid assembly by gap repair in Saccharomyces cerevisiae. *Yeast* **29,** 323–34 (2012).

52. Mikkelsen, M. D. *et al.* Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* **14,** 104–11 (2012).

53. Kildegaard, K. R. *et al.* Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance. *Metab. Eng.* **26,** 57–66 (2014).

54. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–12 (2004).

CLAIMS

1. A eukaryotic cell, such as a yeast cell comprising a bacterial transcriptional activator and a corresponding operator sequence positioned in a eukaryotic promoter, such as positioned within an endogenous promoter of said cell, which activator controls the expression of a gene from said eukaryotic promoter.

2. The cell according to claim 1, wherein the expression of a gene from said eukaryotic promoter is depending on the presence, such as dose dependent, of a ligand specifically binding said transcriptional activator.

3. The cell according to claim 2, wherein said cell comprises a gene encoding the expression of said ligand, one or more genes encoding a pathway of enzymes synthesizing said ligand, and/or a gene encoding a compound that is metabolized into said ligand.

4. The cell according to any one of claims 1-3, wherein said cell comprises an exogenous reporter gene, and/or one or more further regulatory gene, such as a gene encoding antibiotic resistance.

5. The cell according to claim 4, wherein said reporter gene provides for fluorescence output, such as a gene encoding green fluorescent protein, blue fluorescent protein or luciferase.

6. The cell according to any one of claims 2-5, wherein one or more of the genes independently selected from the gene encoding the expression of said ligand, one or more genes encoding a pathway of enzymes synthesizing said ligand, a gene encoding a compound that is metabolized into said ligand, an exogenous reporter gene, and one or more further regulatory gene; is under the control and/or is activated by said eukaryotic promoter.

7. The cell according to any one of claims 1-6, wherein said transcriptional activator is selected from any one selected from table 6, such as any one selected from BenM, FdeR, MdcR, and ArgP.

8. The cell according to any one of claims 2-6, wherein said ligand and transcriptional activator is any one pair selected from table 7, such as any one selected from muconic acid and BenM; Naringenin and FdeR; Malonate and MdcR; and L-arginin and ArgP.

9. The cell according to any one of claims 1-7, which is a yeast cell, such as *Saccharomyces cerevisiae*.

10. The cell according to any one of claim 1-7, which is a mammalian cell, such as a Chinese hamster ovary cell.

11. The cell according to any one of claims 1-9, wherein said promoter is a full length promoter, or a truncated version with upstream activating sequences, such as UAS1 and UAS2 of the CYC promoter, removed.

12. The cell according to claim 10, wherein said promoter is a yeast promoter, such as the full length CYC1 promoter or CYC1 with upstream activating sequences (UAS1 and UAS2) removed.

13. The cell according to claim 10, wherein said promoter is a mammalian promoter, such as the full length CMV promoter.

14. The cell according to any one of claims 1-12, wherein said transcriptional activator work through a promoter-centric mechanism, wherein said transcriptional activator bind to an operator site in said promoter thereby improving its ability to guide RNA polymerase to initiate transcription.

15. The cell according to any one of claims 1, 3-13, wherein said transcriptional activator does not require binding to any other regulatory subunits and/or which cell is without any further engineering or the co-expression of other molecular components regulating said transcriptional activator.

16. The cell according to any one of claims 2-13, wherein said transcriptional activator does not require binding to any other regulatory subunits apart from its specific ligand and/or which cell is without any further engineering or the co-expression of other molecular components regulating said transcriptional activator.

17. The cell according to any one of claims 1-15, which operator sequence is specific for said transcriptional activator within said promoter.

18. The cell according to claims 16, wherein said operator sequence is positioned immediately 6 bp upstream of the TATA box, such as a TATA box 1, such as TATA-1ß, such as anywhere between 6-15 bp, such as anywhere between 6-14 bp, such as anywhere

between 6-13 bp, such as anywhere between 6-12 bp, such as anywhere between 6-11 bp, such as anywhere between 6-10 bp, such as anywhere between 6-9 bp, such as anywhere between 6-8 bp, such as anywhere between 6-7 bp, such as 6 bp upstream of said TATA box of said eukaryotic promoter.

19. The cell according to any one of claims 1-17, wherein said transcriptional activator belongs to the prokaryote super-family of LysR-type transcriptional regulators (LTTRs).

20. The cell according to any one of claims 1-17, wherein said operator is an LTTR operator sequence selected from *BenO, FdeO, MdcO, and ArgO*.

21. The cell according to any one of claims 1-19, wherein said transcriptional activator is co-expressed in said cell, such as from a promoter selected from TEF1, REV1, RNR2 and TDH3.

22. The cell according to any one of claims 1-20, wherein said transcriptional activator is a functional variant with increased activity, such as $BenM^{H110R, F211V, Y286N}$

23. Use of a prokaryotic transcriptional activator as a regulator of transcription in a eukaryotic cell, such as a yeast cell according to any one of claims 1-21, said transcriptional activator being activated by a ligand specifically binding said transcriptional activator to induce the expression of a protein product from a eukaryotic promoter of said cell, said promoter containing the operator sequence corresponding to said transcriptional activator.

24. Use of a prokaryotic transcriptional activator as a metabolite biosensor for measuring the amount of a ligand extracellular of and/or produced by a eukaryotic cell, such as a yeast cell according to any one of claims 1-21, wherein said ligand specifically bind said transcriptional activator to induce expression of a reporter gene from a eukaryotic promoter of said cell, said promoter containing the operator sequence corresponding to said transcriptional activator.

25. Method for measuring the amount of a ligand intracellular or extracellular of a eukaryotic cell, such as a yeast cell; said cell comprising a bacterial transcriptional activator and a corresponding operator sequence, which activator controls the expression of a reporter gene from a eukaryotic promoter of said cell in response to said ligand specifically binding said transcriptional activator; said promoter containing the operator sequence corresponding to said transcriptional activator; said method including the steps of

a) Cultivating a eukaryotic cell as defined in any one of claims 1-21;

b) Measuring the output from said promoter of said reporter gene;

c) Correlating said output from step b) with amount of said ligand.

26. The method according to claim 24, wherein said ligand is not produced by said eukaryotic cell, but is present in a solution of said eukaryotic cell, such as when used to report toxic waste in a soil.

FIG. 1

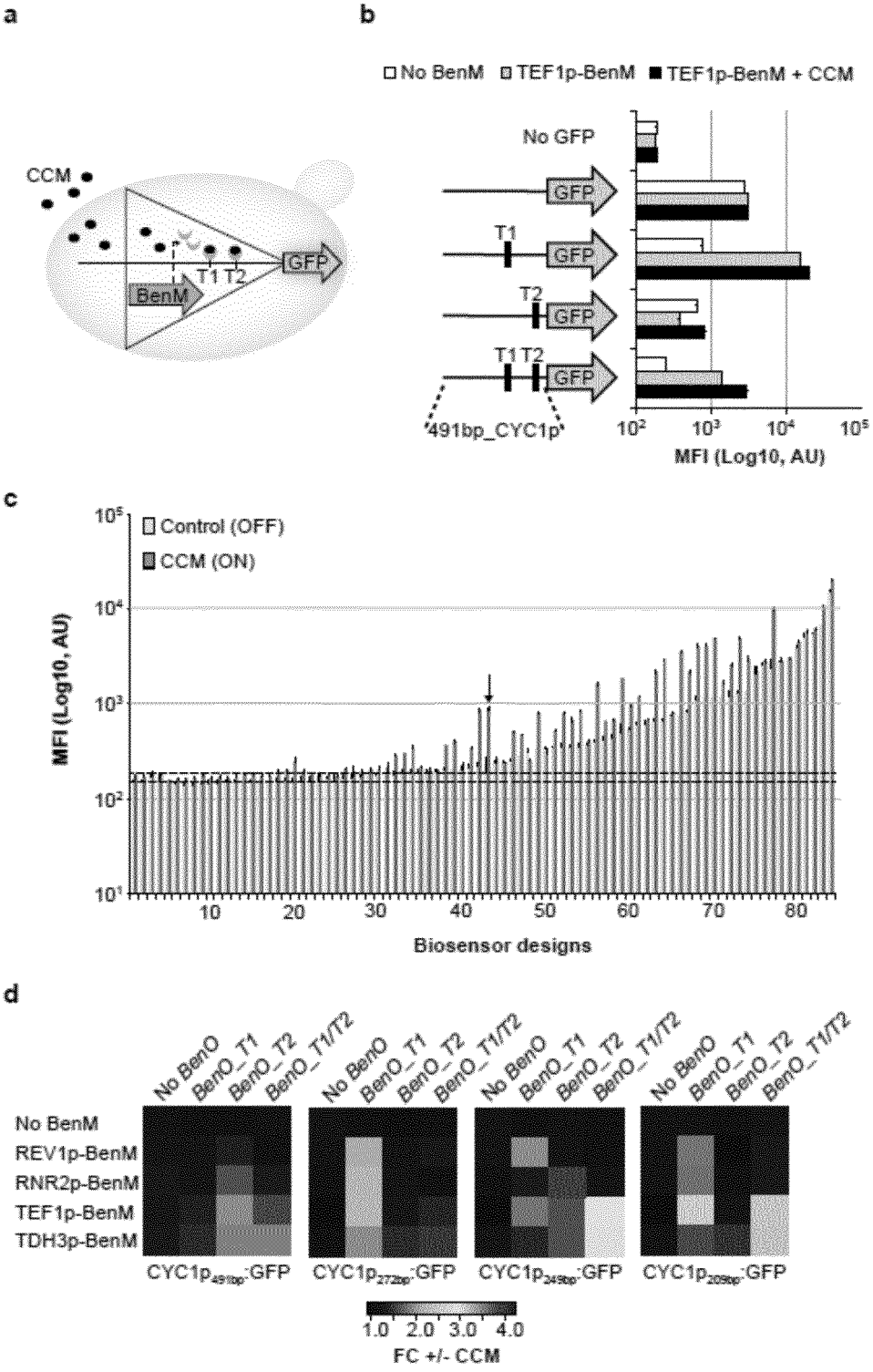

a
b
No BenM
TEF1p-BenM
TEF1p-BenM + CCM
CCM
BenM
T1 T2
GFP
No GFP
T1
T2
491bp_CYC1p
MFI (Log10, AU)
c
Control (OFF)
CCM (ON)
MFI (Log10, AU)
Biosensor designs
d
No BenO
BenO_T1
BenO_T2
BenO_T1/T2
No BenM
REV1p-BenM
RNR2p-BenM
TEF1p-BenM
TDH3p-BenM
CYC1p491bp:GFP
CYC1p272bp:GFP
CYC1p249bp:GFP
CYC1p209bp:GFP
FC +/- CCM

FIG. 2

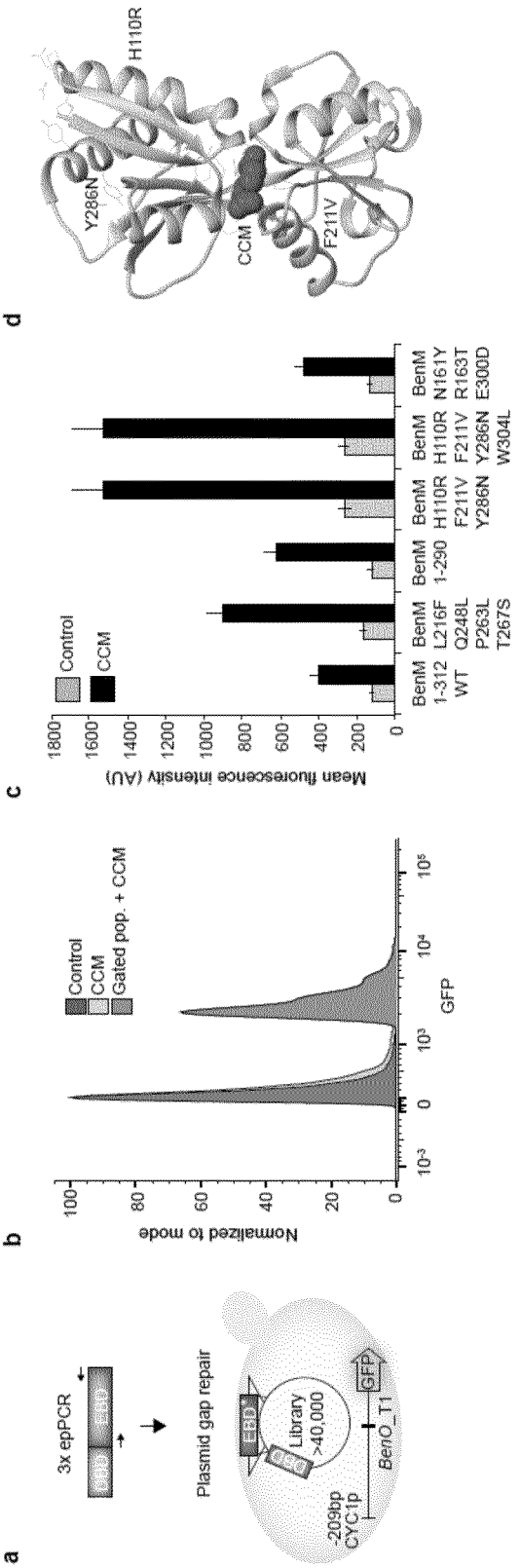
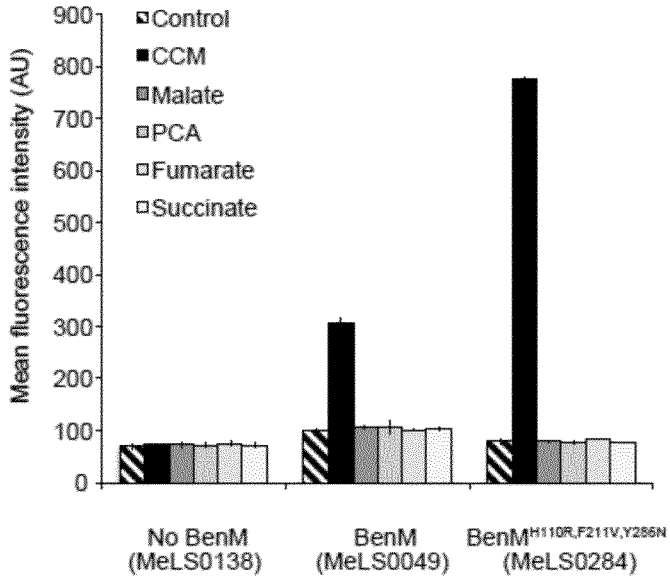

a
3x epPCR
EBD
Plasmid gap repair
Library
>40,000
-209bp
CYC1p
BenO_T1
GFP
b
Normalized to mode
Control
CCM
Gated pop. + CCM
GFP
c
Mean fluorescence intensity (AU)
Control
CCM
BenM 1-312 WT
BenM L216F Q248L P263L T267S
BenM 1-290
BenM H110R F211V Y286N
BenM H110R F211V Y286N W304L
BenM N161Y R163T E300D
d
H110R
Y286N
CCM
F211V

FIG. 3

a


Mean fluorescence intensity (AU)
900
800
700
600
500
400
300
200
100
0
Control
CCM
Malate
PCA
Fumarate
Succinate
No BenM
(MeLS0138)
BenM
(MeLS0049)
BenM H110R,F211V,Y286N
(MeLS0284)

b

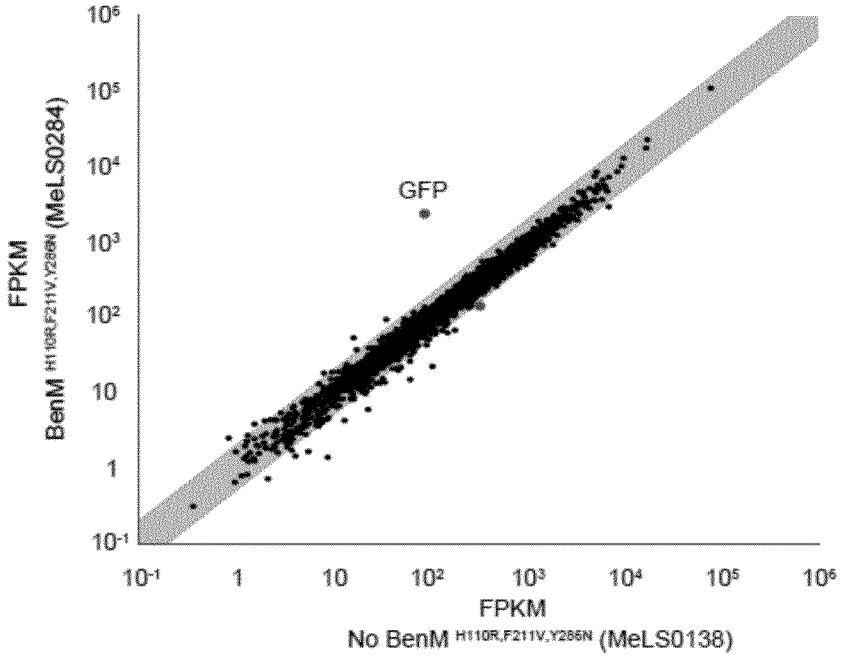

FPKM
BenM H110R,F211V,Y286N (MeLS0284)
10⁶
10⁵
10⁴
10³
10²
10
1
10⁻¹
GFP
10⁻¹
1
10
10²
10³
10⁴
10⁵
10⁶
FPKM
No BenM H110R,F211V,Y286N (MeLS0138)

FIG. 4

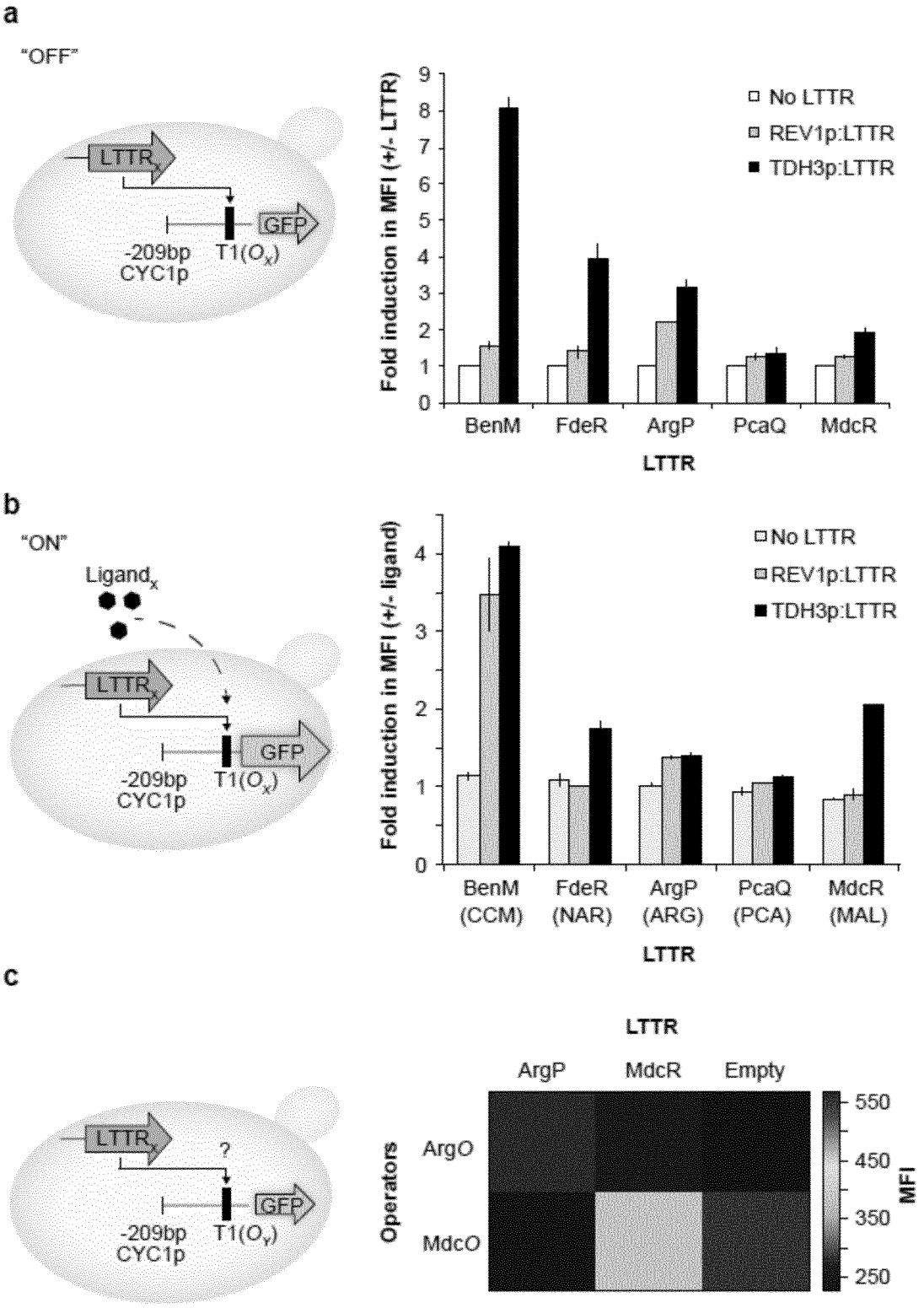

a
"OFF"
LTTR$_x$
GFP
-209bp CYC1p
T1(O_x)
Fold induction in MFI (+/- LTTR)
No LTTR
REV1p:LTTR
TDH3p:LTTR
BenM
FdeR
ArgP
PcaQ
MdcR
LTTR
b
"ON"
Ligand$_x$
LTTR$_x$
GFP
-209bp CYC1p
T1(O_x)
Fold induction in MFI (+/- ligand)
No LTTR
REV1p:LTTR
TDH3p:LTTR
BenM (CCM)
FdeR (NAR)
ArgP (ARG)
PcaQ (PCA)
MdcR (MAL)
LTTR
c
LTTR$_x$
?
GFP
-209bp CYC1p
T1(O_y)
LTTR
ArgP
MdcR
Empty
Operators
ArgO
MdcO
MFI
550
450
350
250

FIG. 5

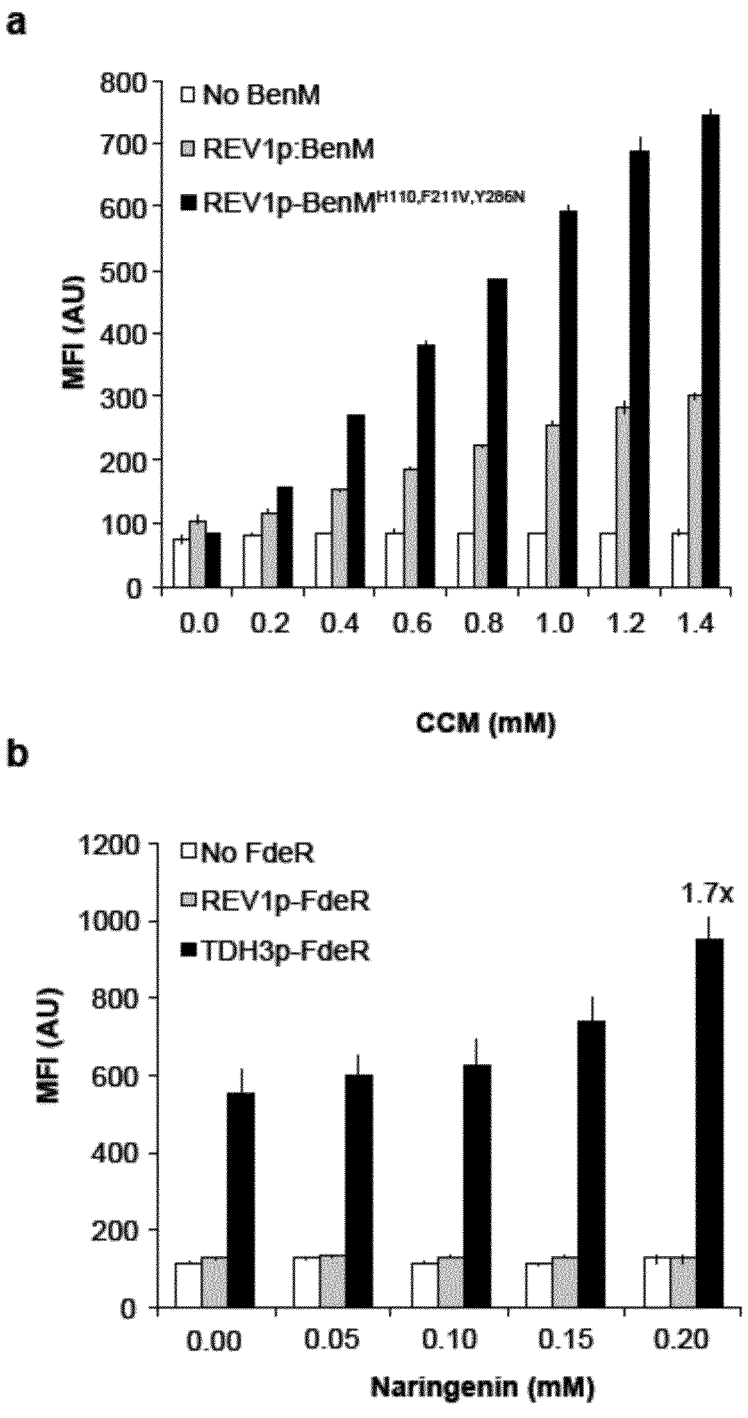

a
800
700
600
500
400
300
200
100
0
MFI (AU)
No BenM
REV1p:BenM
REV1p-BenM H110,F211V,Y286N
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
CCM (mM)
b
1200
1000
800
600
400
200
0
MFI (AU)
No FdeR
REV1p-FdeR
TDH3p-FdeR
1.7x
0.00
0.05
0.10
0.15
0.20
Naringenin (mM)

FIG. 6

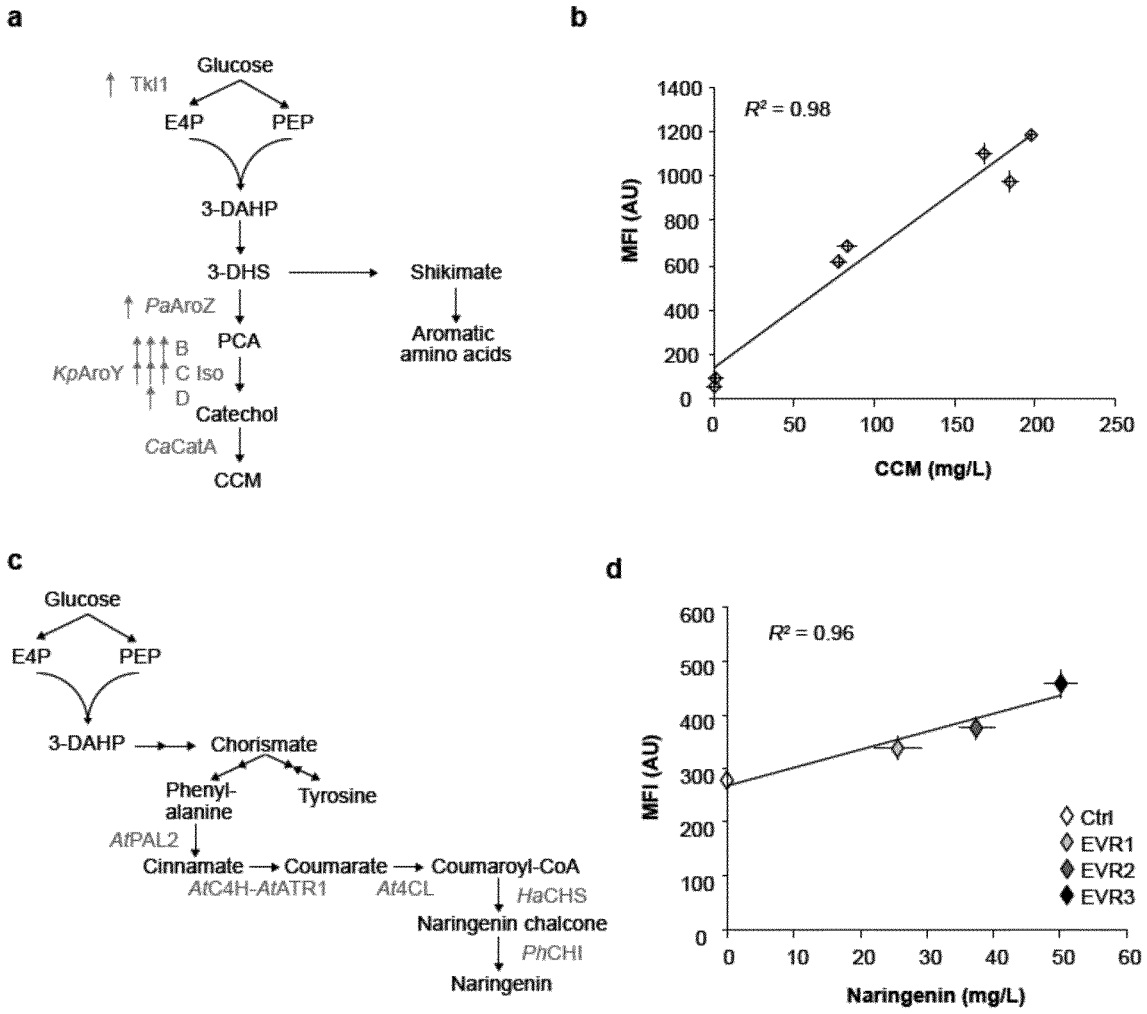

a
Glucose
↑ Tkl1
E4P
PEP
3-DAHP
3-DHS
Shikimate
Aromatic amino acids
↑ PaAroZ
PCA
KpAroY ↑ B ↑ C Iso
↑ D
Catechol
CaCatA
CCM
b
R² = 0.98
MFI (AU)
CCM (mg/L)
c
Glucose
E4P
PEP
3-DAHP
Chorismate
Phenyl-alanine
Tyrosine
AtPAL2
Cinnamate
Coumarate
Coumaroyl-CoA
AtC4H-AtATR1
At4CL
HaCHS
Naringenin chalcone
PhCHI
Naringenin
d
R² = 0.96
MFI (AU)
Naringenin (mg/L)
Ctrl
EVR1
EVR2
EVR3

FIG. 7

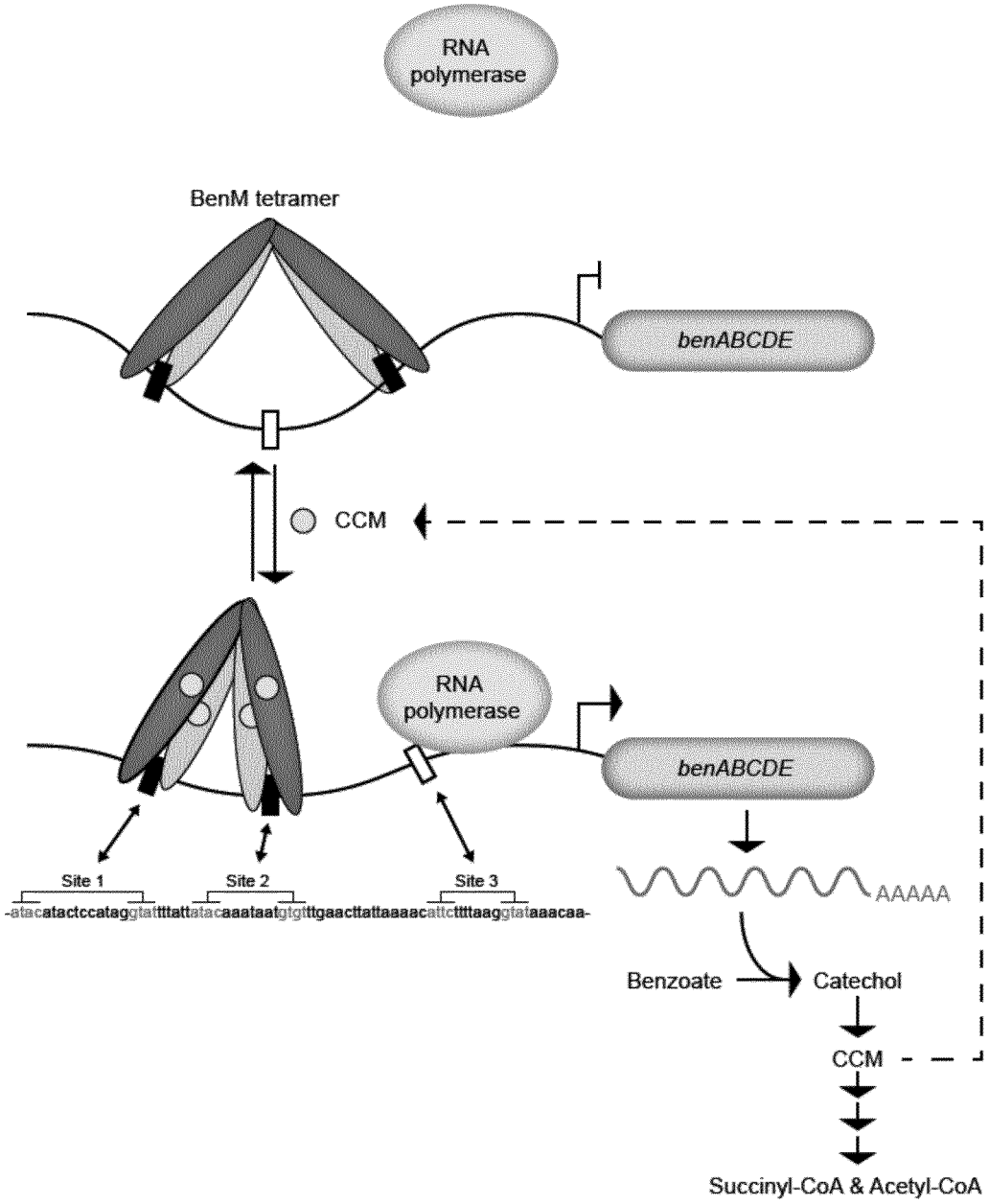

RNA polymerase
BenM tetramer
benABCDE
CCM
RNA polymerase
benABCDE
Site 1
Site 2
Site 3
-atacatactccatagatttttatataaaataatgtgttgaacttattaaaacattctttaaggtataaacaa-
AAAAA
Benzoate
Catechol
CCM
Succinyl-CoA & Acetyl-CoA

FIG. 8

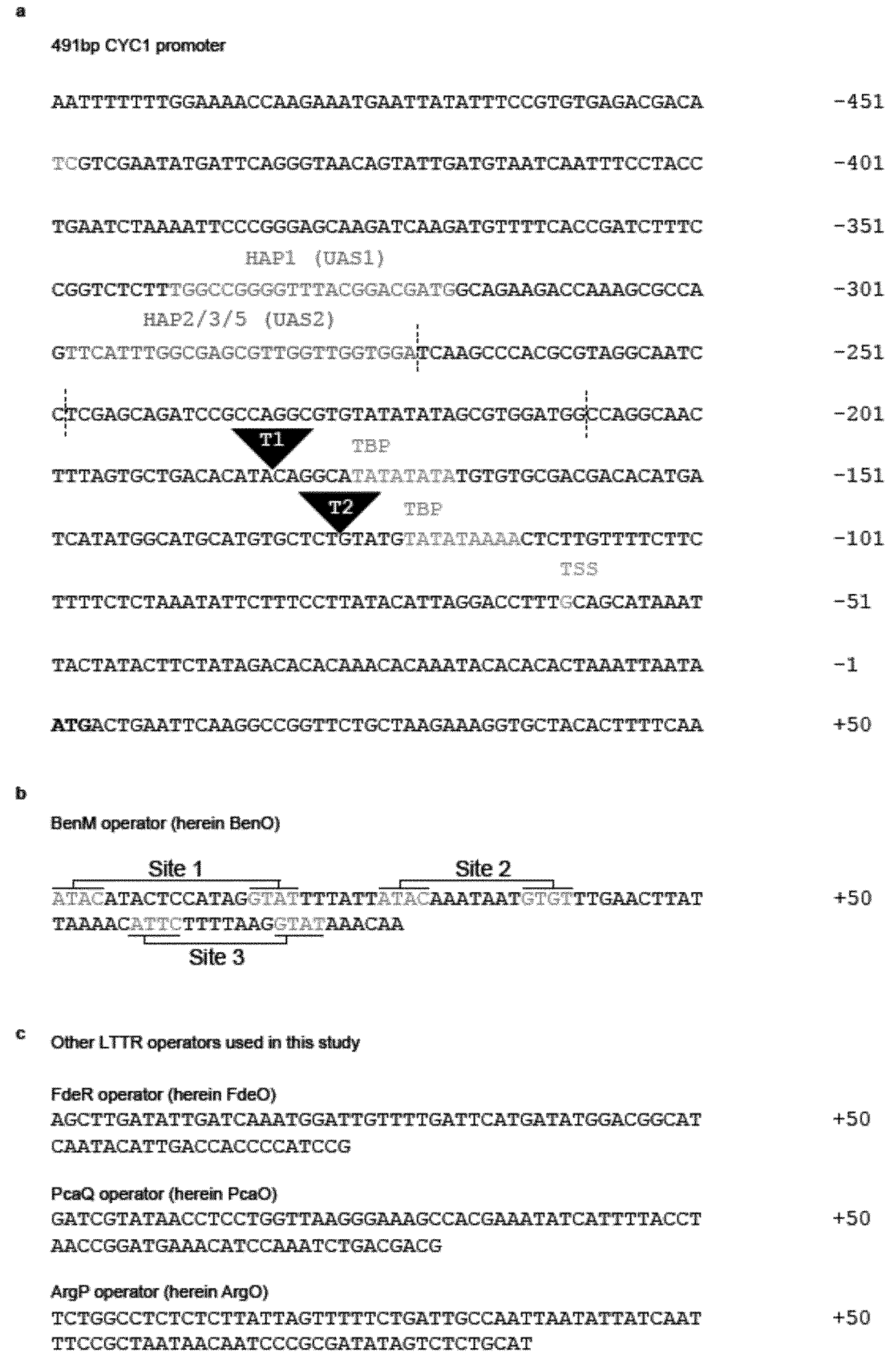

a
491bp CYC1 promoter
AATTTTTTTGGAAAACCAAGAAATGAATTATATTTCCGTGTGAGACGACA -451
TCGTCGAATATGATTCAGGGTAACAGTATTGATGTAATCAATTTCCTACC -401
TGAATCTAAAATTCCCGGGAGCAAGATCAAGATGTTTTCACCGATCTTTC -351
HAP1 (UAS1)
CGGTCTCTTTGGCCGGGGTTTACGGACGATGGCAGAAGACCAAAGCGCCA -301
HAP2/3/5 (UAS2)
GTTCATTTGGCGAGCGTTGGTTGGTGGATCAAGCCCACGCGTAGGCAATC -251
CTCGAGCAGATCCGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAAC -201
T1
TBP
TTTAGTGCTGACACATACAGGCATATATATATGTGTGCGACGACACATGA -151
T2
TBP
TCATATGGCATGCATGTGCTCTGTATGTATATAAAACTCTTGTTTTCTTC -101
TSS
TTTTCTCTAAATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAAT -51
TACTATACTTCTATAGACACACAAACACAAATACACACACTAAATTAATA -1
ATGACTGAATTCAAGGCCGGTTCTGCTAAGAAAGGTGCTACACTTTTCAA +50
b
BenM operator (herein BenO)
Site 1
Site 2
ATACATACTCCATAGGTATTTTATTATACAAATAATGTGTTTGAACTTAT +50
TAAAACATTCTTTTAAGGTATAAACAA
Site 3
c
Other LTTR operators used in this study
FdeR operator (herein FdeO)
AGCTTGATATTGATCAAATGGATTGTTTTGATTCATGATATGGACGGCAT +50
CAATACATTGACCACCCCATCCG
PcaQ operator (herein PcaO)
GATCGTATAACCTCCTGGTTAAGGGAAAGCCACGAAATATCATTTTACCT +50
AACCGGATGAAACATCCAAATCTGACGACG
ArgP operator (herein ArgO)
TCTGGCCTCTCTCTTATTAGTTTTTCTGATTGCCAATTAATATTATCAAT +50
TTCCGCTAATAACAATCCCGCGATATAGTCTCTGCAT

FIG. 9

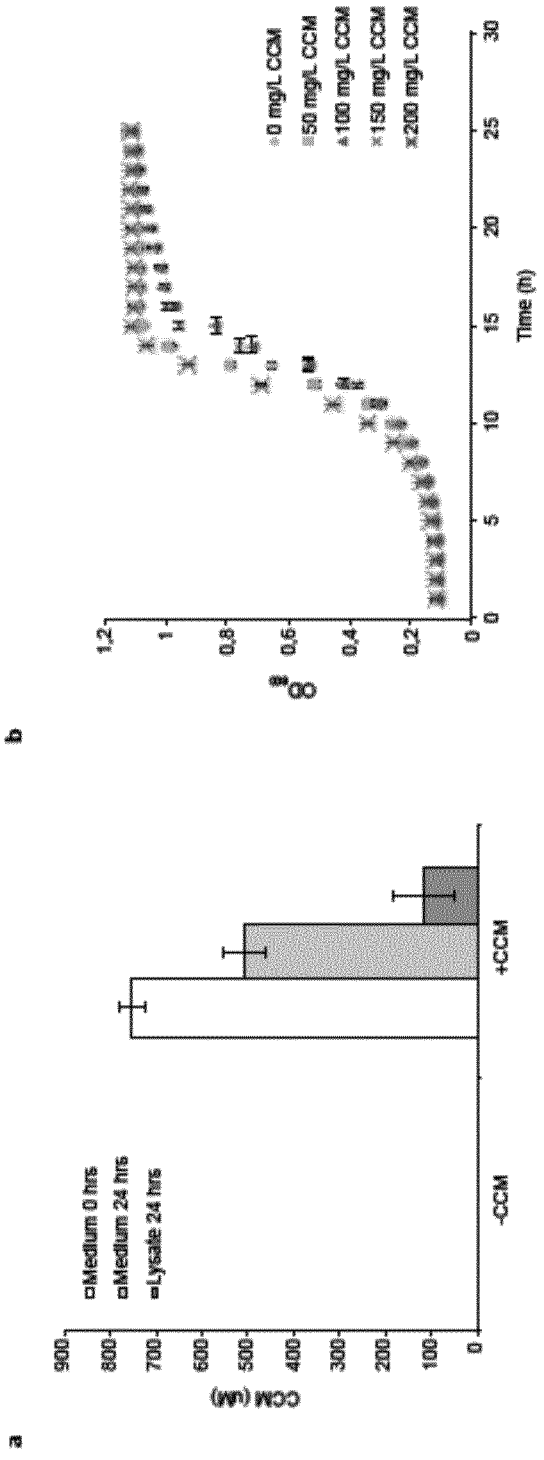

b
0 mg/L CCM
50 mg/L CCM
100 mg/L CCM
150 mg/L CCM
200 mg/L CCM
OD
Time (h)
a
Medium 0 hrs
Medium 24 hrs
Lysate 24 hrs
CCM (uM)
-CCM
+CCM

FIG. 10

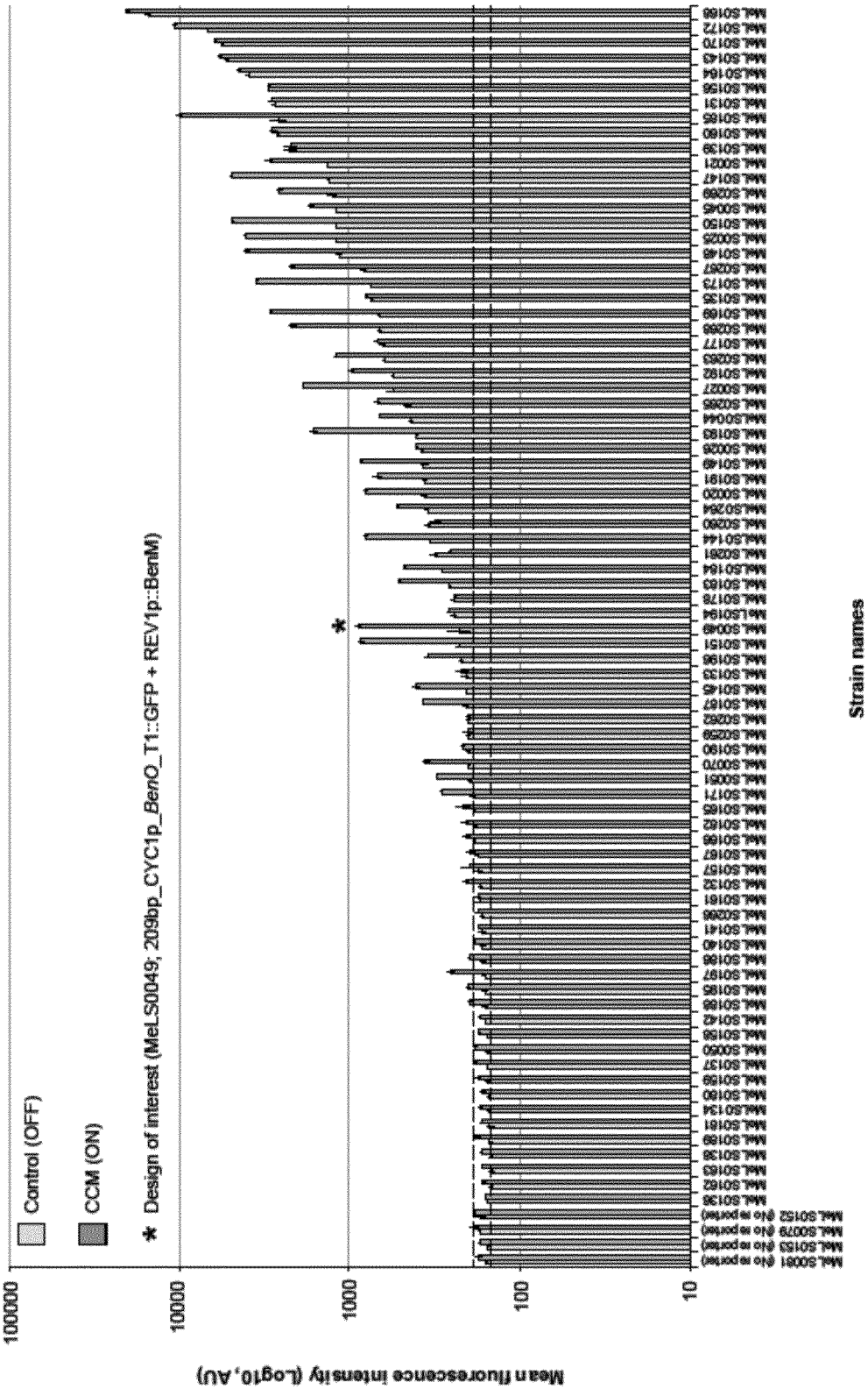
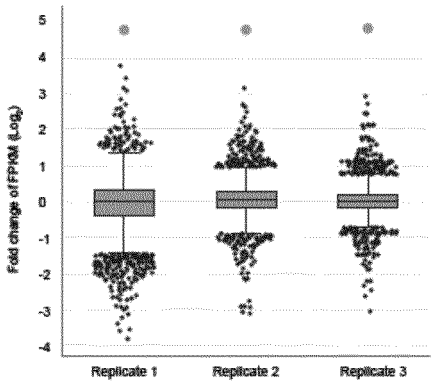

Control (OFF)
CCM (ON)
* Design of interest (MeLS0049; 209bp_CYC1p_BenO_T1::GFP + REV1p::BenM)
Mean fluorescence intensity (Log10, AU)
100000
10000
1000
100
10
Strain names

FIG. 11

a



Parameters	Replicate 1	Replicate 2	Replicate 3
# of genes	5825	5825	5825
Mean (SEM)	-0.068 (0.008)	0.036 (0.006)	0.012 (0.005)
1st percentile	-1.930	-1.028	-0.898
99th percentile	1.461	1.163	1.093

b

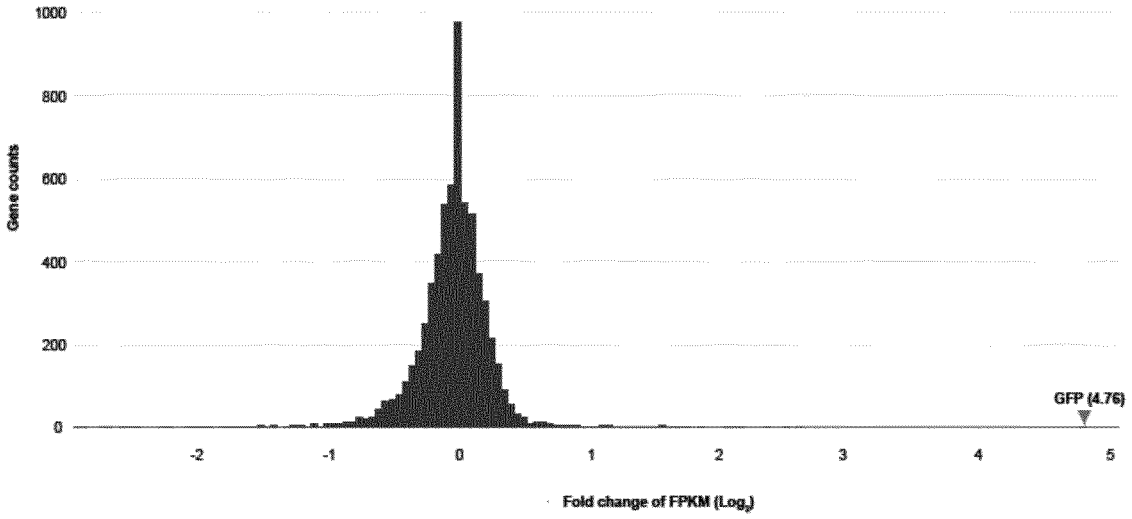


FIG. 12

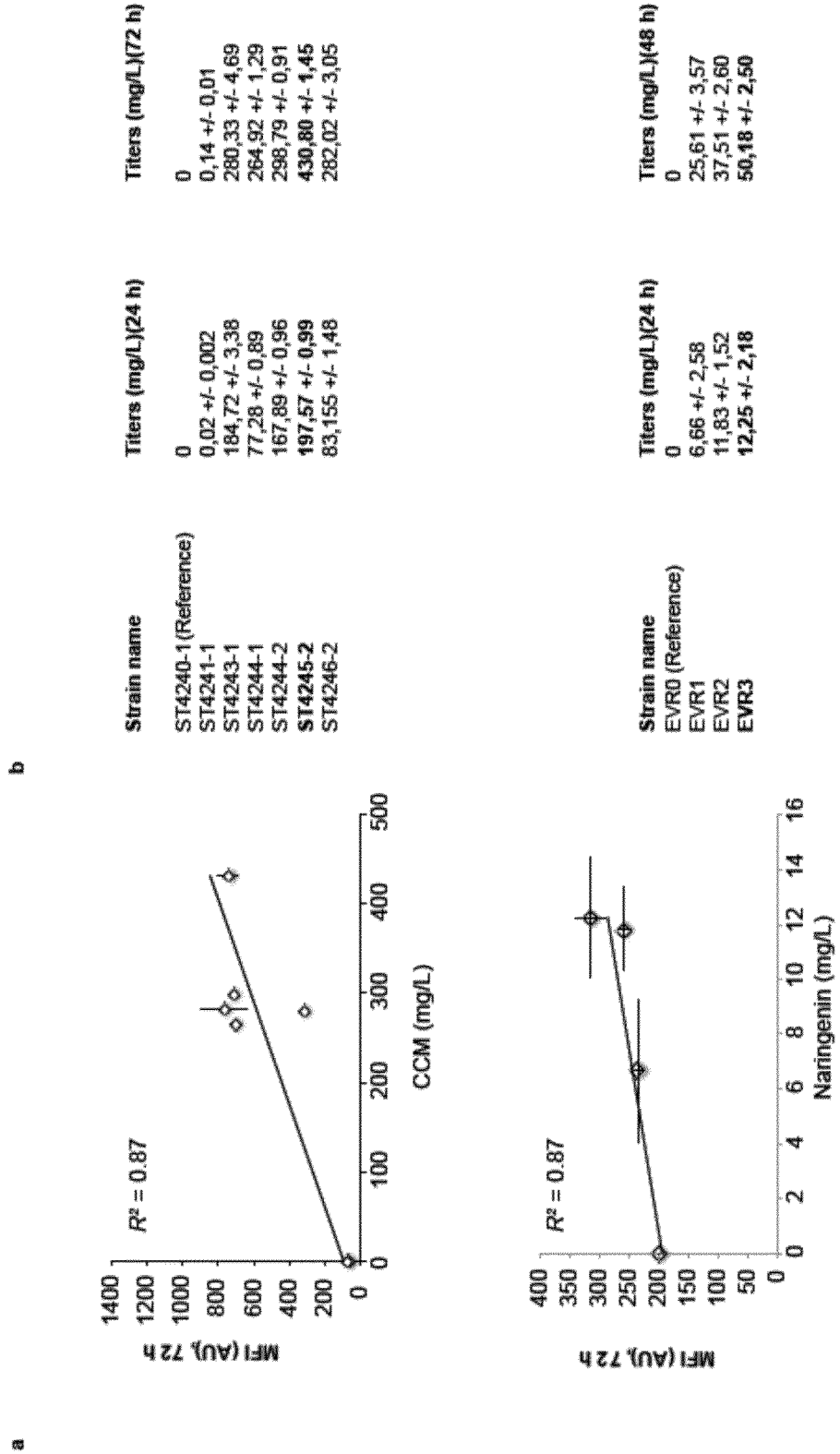

a
MFI (AU), 72 h
R^2 = 0.87
CCM (mg/L)
Naringenin (mg/L)
b
Strain name
Titers (mg/L)(24 h)
Titers (mg/L)(72 h)
ST4240-1 (Reference) 0 0
ST4241-1 0,02 +/- 0,002 0,14 +/- 0,01
ST4243-1 184,72 +/- 3,38 280,33 +/- 4,69
ST4244-1 77,28 +/- 0,89 264,92 +/- 1,29
ST4244-2 167,89 +/- 0,96 298,79 +/- 0,91
ST4245-2 197,57 +/- 0,99 430,80 +/- 1,45
ST4246-2 83,155 +/- 1,48 282,02 +/- 3,05
Strain name
Titers (mg/L)(24 h)
Titers (mg/L)(48 h)
EVR0 (Reference) 0 0
EVR1 6,66 +/- 2,58 25,61 +/- 3,57
EVR2 11,83 +/- 1,52 37,51 +/- 2,60
EVR3 12,25 +/- 2,18 50,18 +/- 2,50

FIG. 13

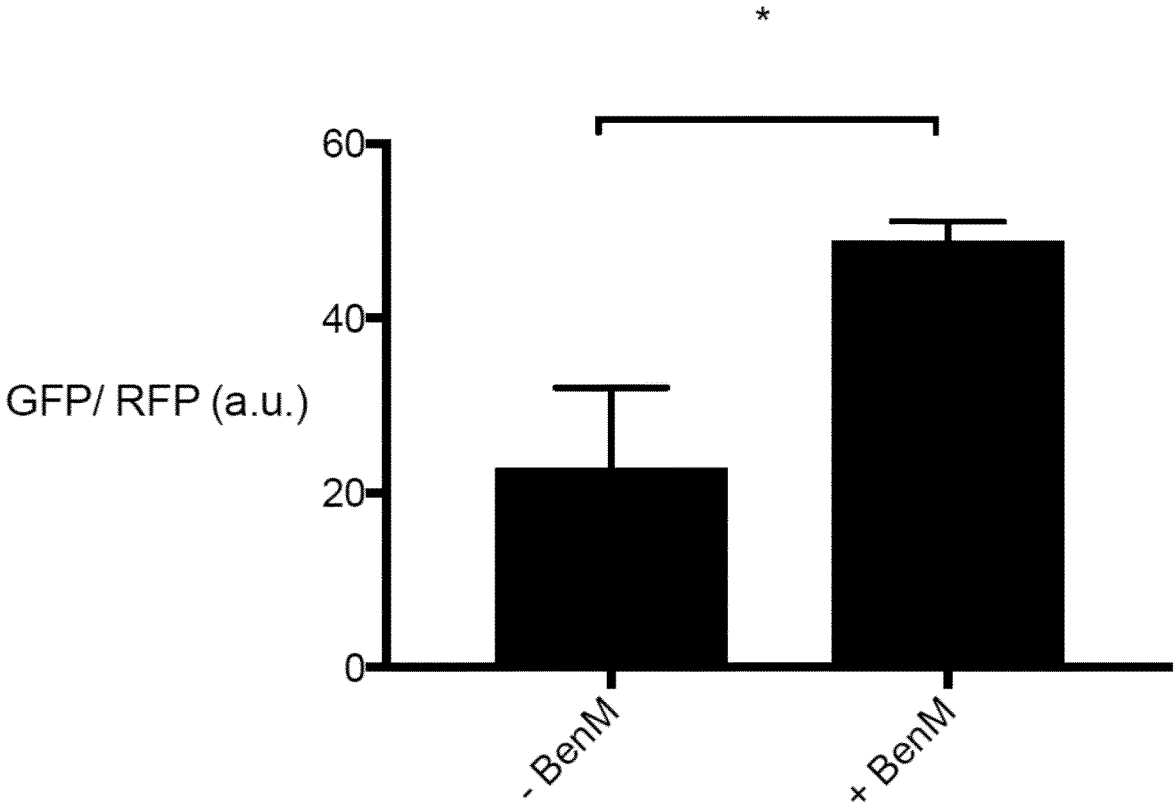
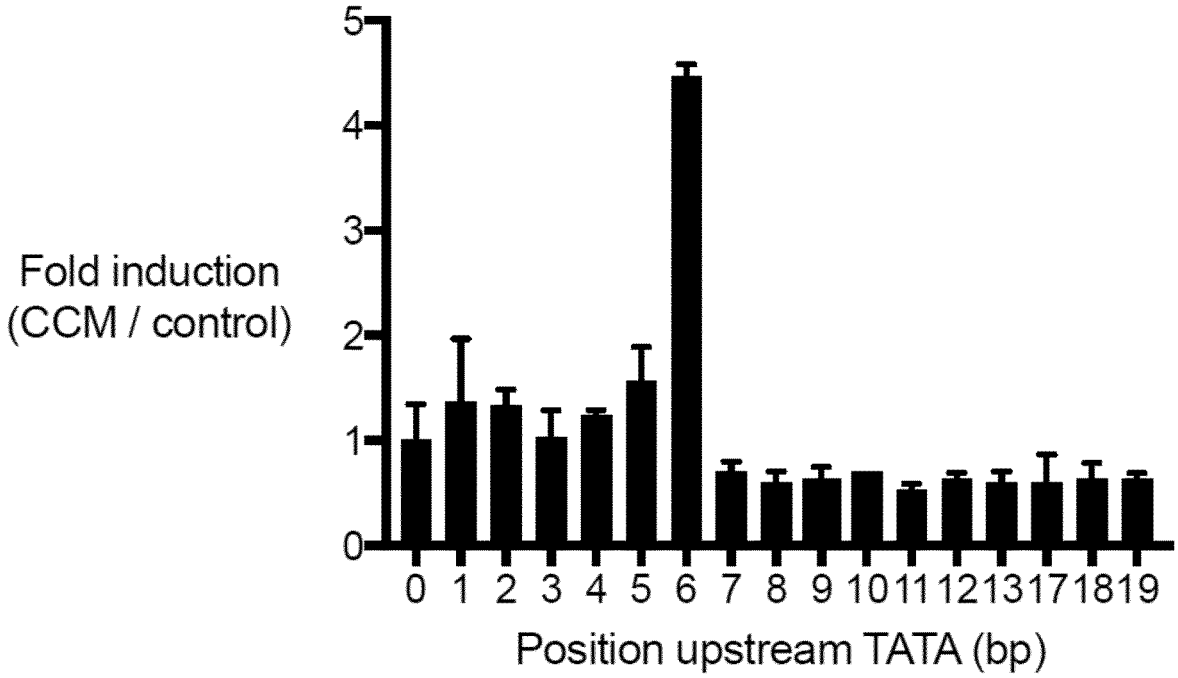


FIG. 14


Fold induction
(CCM / control)
5
4
3
2
1
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 17 18 19
Position upstream TATA (bp)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/066943

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/81 C12N15/67
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, COMPENDEX, Sequence Search, EMBASE, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WEBER ET AL: "Inducible product gene expression technology tailored to bioprocess engineering", CURRENT OPINION IN BIOTECHNOLOGY, LONDON, GB, vol. 18, no. 5, 22 October 2007 (2007-10-22), pages 399-410, XP022350910, ISSN: 0958-1669, DOI: 10.1016/J.COPBIO.2007.09.002 the whole document pages 400-401	1-26

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 September 2017	05/10/2017

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Madruga, Jaime

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/066943

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PETRA NEDDERMANN ET AL: "A novel, inducible, eukaryotic gene expression system based on the quorum-sensing transcription factor TraR", EMBO REPORTS, vol. 4, no. 2, 1 February 2003 (2003-02-01), pages 159-165, XP055375325, GB ISSN: 1469-221X, DOI: 10.1038/sj.embor.embor734 the whole document	1-26
X	WO 96/40946 A1 (UNIV YALE [US]; SCHATZ DAVID G [US]) 19 December 1996 (1996-12-19) the whole document	1-26
X	BRENT R ET AL: "A EUKARYOTIC TRANSCRIPTIONAL ACTIVATOR BEARING THE DNA SPECIFICITY OF A PROKARYOTIC REPRESSOR", CELL, vol. 43, no. 3, PART 02, 1 December 1985 (1985-12-01), pages 729-736, XP009033985, ISSN: 0092-8674, DOI: 10.1016/0092-8674(85)90246-6 the whole document	1-26
A	TSUNG K ET AL: "ENHANCEMENT OF RNA POLYMERASE BINDING TO PROMOTERS BY A TRANSCRIPTIONAL ACTIVATOR, OMPR, IN ESCHERICHIA COLI: ITS POSITIVE AND NEGATIVE EFFECTS ON TRANSCRIPTION", PROCEEDINGS NATIONAL ACADEMY OF SCIENCES PNAS, NATIONAL ACADEMY OF SCIENCES, US, vol. 87, 1 August 1990 (1990-08-01), pages 5940-5944, XP000604714, ISSN: 0027-8424, DOI: 10.1073/PNAS.87.15.5940 the whole document page 5940, column 1	1-26

-/--

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WILLIAM BACCHUS ET AL: "Increasing the dynamic control space of mammalian transcription devices by combinatorial assembly of homologous regulatory elements from different bacterial species", METABOLIC ENGINEERING, vol. 15, 23 November 2012 (2012-11-23), - 23 November 2012 (2012-11-23), pages 144-150, XP055375308, US ISSN: 1096-7176, DOI: 10.1016/j.ymben.2012.11.003 page 144, column 2 page 146, column 2 -----	1-26
X	PTASHNE M: "HOW EUKARYOTIC TRANSCRIPTIONAL ACTIVATORS WORK", NATURE, NATURE PUBLISHING GROUP, UNITED KINGDOM, vol. 335, no. 6192, 20 October 1988 (1988-10-20), pages 683-689, XP000142836, ISSN: 0028-0836, DOI: 10.1038/335683A0 the whole document -----	1-26
A	D. TROPEL ET AL: "Bacterial Transcriptional Regulators for Degradation Pathways of Aromatic Compounds", MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, vol. 68, no. 3, 1 September 2004 (2004-09-01), pages 474-500, XP055040116, ISSN: 1092-2172, DOI: 10.1128/MMBR.68.3.474-500.2004 the whole document -----	1-26
A	SARAH E MADDOCKS ET AL: "Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins", MICROBIOLOGY, SOCIETY FOR GENERAL MICROBIOLOGY, GB , vol. 154, no. 12 1 December 2008 (2008-12-01), pages 3609-3623, XP002698452, ISSN: 1350-0872, DOI: 10.1099/MIC.0.2008/022772-0 Retrieved from the Internet: URL:http://mic.sgmjournals.org/content/154/12/3609 cited in the application the whole document ----- -/--	1-26

1

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/066943

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	METTE L SKJOEDT ET AL: "Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast", NATURE CHEMICAL BIOLOGY, vol. 12, no. 11, 19 September 2016 (2016-09-19), pages 951-958, XP055375361, Basingstoke ISSN: 1552-4450, DOI: 10.1038/nchembio.2177 the whole document	1-26
T	& Mette L Skjoedt ET AL: "Supplementaty Information: Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast", Nature Chemical Biology, 19 September 2016 (2016-09-19), pages 951-958, XP055375611, DOI: 10.1038/nchembio.2177 Retrieved from the Internet: URL:http://www.nature.com/nchembio/journal/v12/n11/extref/nchembio.2177-S1.pdf [retrieved on 2017-05-23] the whole document	1-26

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2017/066943

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9640946 A1	19-12-1996	AT 321871 T	15-04-2006
		AU 6274596 A	30-12-1996
		DE 69635992 T2	01-02-2007
		EP 0832254 A1	01-04-1998
		JP H11507539 A	06-07-1999
		US 5851796 A	22-12-1998
		WO 9640946 A1	19-12-1996